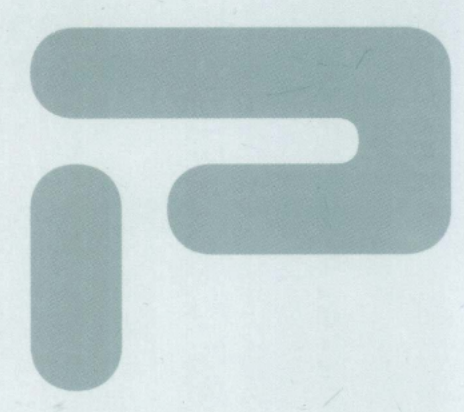
万卷方法 | 心理学研究方法丛书
中国心理学会心理学教学工作委员会推荐

DISSERTATIONS AND THESES FROM START TO FINISH: PSYCHOLOGY AND RELATED FIELDS



学位论文全程指南：心理学及相关领域

约翰·科恩(John D.Cone)
夏侬·福斯特(Sharon L.Foster) 著

张 明等 译



重庆大学出版社
http://www.cqup.com.cn

YUEWEI LUNWEN QUANCHENG ZHINAN:XINLIXUE JI XIANGGUANLINGYU

本书集两位教授20余年来指导学位论文之经验，科学、系统、全面、贴心地为学生提供指导和帮助。

本书提倡将学位论文作为一个工作项目来完成。完成学位论文不仅仅是一个写作过程，而是一个需要多方准备的、环环相扣的系统工程。

首先介绍了“学位论文”是什么，以使学生在一开始就明确他们的工作目标。然后分别介绍如何评估论文工作准备、时间与问题管理、形成和交流研究计划、查找文献形成综述、选择恰当的测量工具获得数据、选择合适的统计方法处理数据、报告研究结果、论文开题与答辩、研究发表等各个环节中的工作步骤和注意事项。

介绍详细而全面，讨论的问题具有代表性。个别环节美国与中国情况不太一样，但这并不影响正在和即将开始准备学位论文的学生从书中受益。

本书所举范例多与心理学相关，但完成论文的流程、技巧与经验等则是多个学科通用的。本书可为社会科学各学科的研究生及导师们提供参考。

参阅及发表与本书相关的评论,请登录

http://q.blog.sina.com.cn/fafang (万卷方法与学术规范博客圈)

上架建议：学术社科

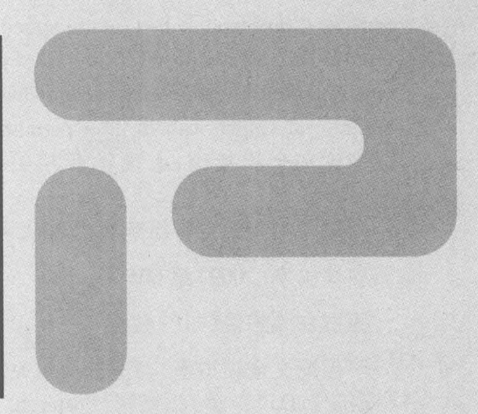
ISBN 978-7-5624-6113-5



定价: 34.00元

万卷方法 | 心理学研究方法丛书
中国心理学会心理学

DISSERTATIONS AND
THESES FROM START TO FINISH:
PSYCHOLOGY AND
RELATED FIELDS



学位论文全程指南：
心理学及相关领域

约翰·科恩 (John D.Cone)
夏侬·福斯特 (Sharon L.Foster) 著
张 明等 译

重庆大学出版社

图书在版编目(CIP)数据

学位论文全程指南:心理学及相关领域/(美)科恩(Cone, J. D.),(美)福斯特(Foster, S. L.)著;张明等译. —重庆:重庆大学出版社,2011.8
(万卷方法)
书名原文:Dissertations And Theses from Start to Finish: Psychology And Related Fields
ISBN 978-7-5624-6113-5

Ⅰ.①学… Ⅱ.①科…②福…③张… Ⅲ.①心理学—学位论文—写作 Ⅳ.①G642.477

中国版本图书馆 CIP 数据核字(2011)第 061491号

学位论文全程指南:心理学及相关领域

约翰·科恩 (John D. Cone)
夏侬·福斯特 (Sharon L. Foster) 著
张 明 等 译
策划编辑:雷少波 林佳木
责任编辑:杨 敬 版式设计:林佳木
责任校对:任卓惠 责任印制:赵 晟

*

重庆大学出版社出版发行
出版人:邓晓益
社址:重庆市沙坪坝正街 174 号重庆大学(A 区)内
邮编:400030
电话:(023) 65102378 65105781
传真:(023) 65103686 65105565
网址:http://www.cqup.com.cn
邮箱:fxk@cqup.com.cn (营销中心)
全国新华书店经销
重庆升光电力印务有限公司印刷

*

开本:787×1092 1/16 印张:15 字数:336千
2011 年 8 月第 1 版 2011 年 8 月第 1 次印刷
印数:1—3 000
ISBN 978-7-5624-6113-5 定价:34.00 元

万卷方法是重庆大学出版社自2004年7月以来陆续出版的一套深入、系统地介绍社会科学研究方法问题的书系，至今已出版图书84个品种。

“万卷方法”的理想

为社会学、政治学、管理学、经济学、传播学、心理学、教育学等社会科学各领域的学者和研究生提供一个内容规范、使用便捷的“研究方法工具箱”。

“万卷方法”的受众

- 图书馆及大学社会科学各院系的资料室。
- 社会科学各领域的研究人员。
- 社会科学各专业的研究生和本科生。
- 社会及市场调查的实务工作者。

欢迎致信 wjffsyh@foxmail.com 加入万卷方法书友会，详情可登录以下网址了解。

万卷方法与学术规范博客圈：
http://q.blog.sina.com.cn/fafang

万卷方法——心理学研究方法丛书
选编委员会

作者简介

约翰·科恩(John D. Cone)博士

获得斯坦福大学心理学专业学士学位、华盛顿大学硕士和博士学位。曾就职于普吉特湾大学、西弗吉尼亚大学、夏威夷大学、美国国际大学、州立圣迭哥大学、联盟国际大学。美国心理学协会和美国心理学会会员,行为分析学会会员,执业行为分析师。研究兴趣主要包括个体特质发展评估方法,孤独症干预研究,发展、实施和评估针对发展障碍个体的大规模服务系统等。作为高级主管和组织顾问,他出版了多部著作,包括《评估产出:有效实践的经验主义工具》(*Evaluating Outcomes: Empirical Tools for Effective Practice*,2001)等。在业余时间里,他喜欢跑步、打理渔船,以及在蔚蓝的太平洋上航行。

夏依·福斯特(Sharon L. Foster)博士

加利福尼亚州圣迭哥市联盟国际大学著名学者。1978 年于华盛顿大学医学院实习结束后,获得纽约州立大学斯托尼布鲁克分校心理学博士学位。曾就职于西弗吉尼亚大学。任《行为评价》(*Behavioral Assessment*)和《咨询与临床心理学》(*Journal of Consulting and Clinical Psychslogy*)杂志的副主编,行为科学高级研究中心成员(2000—2001)。研究主要集中在儿童朋辈关系,亲子冲突的评估与应对,以及研究方法等方面。发表了大量研究论文,参与编纂多部书籍,并出版了 4 部著作。

总 序

自古以来，人类一直在探索着自己的内心世界：我到底是个什么样的人？为什么在许多方面我与周围人们如此相似，而在其他方面又如此不同？我们是怎样认知世界的？为什么有时记忆会错误有时记忆会很牢固？喜怒哀乐、爱恋、责任感是如何产生的？情绪能自我调控吗？意识是怎么回事？梦是怎么回事，它能预测未来吗？为什么一个人独处时和在群体中的行为是不一样的？我们是怎样理解语言的，又是怎样组织和表达语言的？如此等等的问题，从传说、神话、甲骨文中我们都能发现，可见，人类长期以来一直对理解自身的内心世界有着浓厚的兴趣。然而直至1879年冯特(Wilhelm Wundt, 1832—1920)在莱比锡大学建立了心理实验室之后，人类才开始借助科学方法来寻求这些问题的答案。通过精密严格的数据收集与事实分析来研究心理与行为，积累知识，从而发展出今日的心理学。而随着心理学的发展，其研究方法也发展起来。

心理学家的科学研究是一种自觉的、有目的地探索精神世界的求知活动。这种探索，不仅有理论，还有与理论有关的观点及方法、仪器等。当代心理学主要有五种理论取向在探讨人类内心世界的奥秘，并对大脑如何工作提出了不同的假设，因而它们所采用的方法也不同。持生物学理论取向的心理学家认为心理是脑的机能，采用脑电图(EEG)、正电子发射断层扫描技术(PET)、功能核磁共振成像技术(FMRI)等方法来探讨人类是如何产生知觉、记忆、推理、情绪和某些人格特征的。持学习理论取向的心理学家认为心理是个体对环境条件作用积累经验的生理变化，采用操作条件作用、奖赏、惩罚、观察学习等方法来探讨人类和动物行为的形成及矫正。持认知理论取向的心理学家把人类的心理活动视为类似于计算机的信息加工，用反应时、正误率和口语报告等手段来探讨人类的知觉、记忆、言语和思维等心理过程。持精神分析论取向的心理学家把人的心理视为潜意识本能的表现，采用个案调查、诠释学方法来探讨个人内部的驱力、冲突或心理疾病等潜意识活动。持人本主义理论取向的心理学家把人的心理视为自我实现需求的表现，采用相关法、诠释学方法来探讨自由意志、个人成长、潜能实现等问题。在心理学研究中理论观点与方法始终是结合在一起，相辅相成的。它们既是指导这种探索活动的武器，又是保证这种活动取得成果的基础。正因为有了这一套套理论观点与方法的有机结合，心理学的科学研究才成为一种自觉的、有目的的定向活动，心理学也才成其为科学。因此，学习心理学研究方法具有十分重要的意义。

首先，有助于我们自觉地将理论与方法相结合，养成科学思维的习惯。心理学家的科学研究都具有明确的目的，即需要解决的问题，为此，就要对该问题以往的研究和目前的现状进行文献综述，并按照一定的有效程序对其进行探讨。即是说，心

理学研究的基本程序与任何科学研究是一样的,都包含下列步骤:选题和提出假设——设计研究方案(用以检验假设的真伪)——收集资料——整理和分析资料——解释结果和检验假设。从心理学研究程序的各个环节可以看出,在心理学研究中,精密的仪器和先进的实验设备固然重要,然而最重要的还是研究者的头脑。通过对心理学研究方法的学习,将有助于我们养成善思考和科学思维的习惯。心理学研究方法,不仅可以帮助我们运用自己的智慧去进行科学研究,而且它还可以帮助我们去鉴别自己和他人的研究成果的正确与谬误。正因为如此,有经验的学者在评价一篇学术论文时往往不只是看它的结论,而且,甚至是更重要的还要看论文作者是通过怎样的途径和方法而获得结论的。

其次,有助于激发我们的创新观念和达成创新目标。心理学的理论、观点和研究方法是多样性的,其研究成果的科学性也不同。学习了心理学研究方法之后,我们了解到心理学研究的各种方法,如个案法、相关法和实验法在心理学研究的不同时期有不同的用途,其信效度也是不同的。个案研究在心理学研究初期是有用的(有助于发现可供研究的现象和变量),但是要确定变量之间的因果关系,建立科学理论,则必须借助于实验法。弗洛伊德(Sigmund Freud, 1856—1939)根据自己对歇斯底里病人的临床观察和对梦、失误和笑话等的现象分析,建立起以潜意识动机为基础的精神分析理论。这个理论不是一个科学的理论,其中许多概念和命题缺乏实证效度。然而,熟悉实验法的学者想到用实验法来检验弗氏理论中的许多概念。例如,对于潜意识这种现象,他们发展出一系列的内隐实验程序来进行检验,结果发现有内隐学习、内隐记忆、内隐情绪、内隐动机的存在。虽然 20 世纪 80 年代兴起的心理潜意识研究与弗洛伊德的潜意识的性本能和死亡本能有本质的区别,但却加深了我们对心理成分和潜意识性质的认识。自我是心理学研究的一个主题,在科学心理学的早期,一些著作从理论上探讨自我的性质,到了 20 世纪 70 年代不少学者开始用相关法和实验法探讨自我的成分和机能,后来由于引入神经科学方法(如脑的电刺激、功能核磁共振成像技术)才开始探索自我的脑机制。正如巴甫洛夫(И. П. Павлов, 1849—1936)所说"科学是随着研究方法所获得的成就而前进的。研究方法每前进一步,我们就更提高一步。随之在我们面前也就开拓了一个充满着种种新的、更加广阔的远景。因此,我们头等重要的任务乃是制定研究方法。"①

第三,有助于年轻心理学工作者快速成长。做任何事情都要讲究方法。方法对头,事半功倍;方法不对,事倍功半,甚至导致失败。心理学史表明,有些心理学家之所以能在学科上有所建树、有所贡献,除了他们的天赋聪慧,当时的科技水平和良好的学术环境外,还往往与他们能正确运用新的研究方法有密切的关系。系统地学习心理学研究方法显然比只凭个人经验、漫无边际地去摸索,更能促进年轻心理学工作者的快速成长。

2005 年秋天,重庆大学出版社"万卷方法"总策划雷少波同志带着已出版的新书来征求我对翻译这套图书的意见。我很高兴看到他们对这套书的设想:"万卷方法"是重庆大学出版社从 2004 年开始出版,拟系统深入地介绍各门社会科学研究方法的大型工具性丛书,其中包括心理学研究方法。这是一项促进我国心理学事业发

① 巴甫洛夫选集[M].北京:科学出版社,1955:49.

展的开创性工作，我给予热情鼓励和支持。在我看来，“万卷方法——心理学研究方法丛书”具有以下特点：

(1)品位高。对于研究方法的著作来说，质量优、品位高是最重要的。丛书所介绍的是国外心理学领域中，许多有成就的心理学家所普遍认可的心理学研究工作的原理、方法，研究课题设计，以及如何正确使用各种可以应用的技术手段等。例如如何做心理学实验、如何进行心理学的质性研究、如何撰写心理学学术论文，以及在心理研究中如何使用数理统计、应遵循哪些伦理道德等。其中像心理学质性研究、心理学研究伦理道德等著作，国内至今未见有专题著作翻译出版，是国内急需的。

(2)适用面广。丛书所介绍的心理学研究方法是相对基础性的，可供高校心理学专业的本科生和研究生作为教材或教学参考书使用，也可供广大人文社会科学工作者参考。

(3)开放性。根据我国心理学教学和科研的需要以及心理学研究方法的发展，出版社将通过版权引进和本土开发，使丛书不断丰富与完善。

我相信广大读者会喜爱“万卷方法——心理学研究方法丛书”，祝愿“万卷方法”不断发展，日益完善。

是为序。

黄希庭

谨识于2009年10月

西南大学 有容斋

译者序

诗人泰戈尔在《世上最远的距离》中曾写道,“世上最远的距离/不是生与死/而是我站在你面前/你却不知道我爱你”。后来,一些有志于科学研究的年轻人又发现,原来世上最遥远的距离并不是爱人站在面前自己却全然不知,而是论文标题与参考文献之间的距离。从标题到参考文献之间短短几页或几十页的间隔,怎么就成了世间最遥远的“距离”了呢?这固然是人们在学术研究之余聊以自嘲的玩笑话,但也反映出一些刚刚开始学术研究的学生在进行论文撰写过程中遇到的困难和困惑。怎样选择合适的题目?如何回顾以往的研究成果?怎样分析、报告自己的研究结果并对其加以讨论?如何将自己的论文发表出来接受同行的检阅?这些问题归结起来,其实就是如何缩短从开始着手进行论文撰写到最终完成写作之间的距离的问题。

科恩教授和福斯特教授的这本《学位论文全程指南:心理学及相关领域》正是旨在帮助我们缩短“世上最远的距离”。借由两位教授多年的研究写作经历和经验,本书为年轻研究者提供很多具体可操作且有益的指导意见。相信利用前人摸索出来的经验,能使年轻研究者少走弯路、少浪费时间,更快地步入正轨。但同时需要强调的是,作者的指导意见只是他自己经验的总结,切不可全盘照搬,或仅仅是抱住此书进行理论学习并幻想着一蹴而就。只有借鉴书中提到的方法,并不断在真实的论文写作过程中进行操练,才能达到拉近论文标题和参考文献之间“距离”的目的,更快更好地写出优质的学位论文。

这本书由我主持翻译,参与翻译的有东北师范大学心理学专业的温婷(第1—5章)、张天阳(第6章)、陈丽红(第7—9章)、王佳莹(第10—14章)同学;此外,2006级心理学专业的王好博、栾子烟、董波、吴丽娇、陈瀛等同学在书稿的编辑和审校过程中也做了大量的工作。借此机会对参与书稿翻译工作的同学致以由衷的感谢,正是各位辛勤的劳动才使这本书的中文版得以与读者见面。

最后,全书由我进行了统校。尽管我和各位译者花费了很多的时间和精力在此书稿之中,但译本中的错误在所难免,还望读者诸君不吝指正。

张　明

东北师范大学心理学系　教授

二〇一一年四月于长春

作者前言

15 年前我们曾编写了《学位论文全程指南:心理学及相关领域》一书,当时用以指导学生进行论文研究和写作的指导手册还很少。本书的第 1 版出版后,我们只是尝试性地对内容作了一些调整。毕竟,当时研究的一些基本准则并未发生改变:就是既不同于总统竞选也不同于言语辩论中所需要的良好表达技能。然而,随着时间推移,这一领域中的内容发生了很多变化,出现了一些新的研究伦理道德、论文发表手册以及统计检验方法。此外,在行为科学研究领域中也出现了一些关于方法论的新思想,尤其是信息革命以及个人计算机和网络的普及,使得学生和研究者资料收集的方式和途径、数据管理和分析所使用的工具相应地发生了转变。这些转变促使我们对本书的内容进行调整。

本书第 1 版的主要目的是要给学生论文写作提供帮助。多年来,我们收到了许多来自学生和指导教师的信件、邮件或其他方式的交流和反馈,在此表示感谢! 正是这些反馈让我们相信本书实现了最初的编写目的。同时,我们也收到了很多建设性的批评建议,并且认识到书中的一些内容已经无法满足研究发展的要求。在这次修订中,我们对每一章内容都进行了必要的更新,根据一些同事的建议,舍弃了一些落后于当前发展水平的内容和资料,增加了一些旧版中没有介绍的新内容。在修订过程中,我们力求使书的内容与 21 世纪技术发展水平相一致,书中有关论文研究过程中的文献综述、准备工作、统计和后期数据分析这几章的内容最大限度地体现了这种更新。在每章的最后我们列出了一些新的延伸性阅读文献和资源,读者可以通过这种途径收集更多的信息资料。此外,我们也列出了一些可以提供有价值资源和信息的临时性网址,在不久的将来这些网址可能会长期存在和使用。对于这些网络资源我们持一种谨慎的态度,虽然网络资源的获得相对于专业书籍和期刊文献更为快捷迅速,但是通过网络获得的资源却不如纸质出版物容易通过同行审核。

在新版中,我们保留了一些学生反馈中提到的有价值的内容。我们坚持,并且美国心理学会已经同意本书以学生能够负担的合理价格以平装形式出版。与本书第 1 版的编写目的一样,我们希望这一版能为读者在研究过程中提供有价值的思想和指导,让他们从研究最初的疑惑走向激动和最后的满意,让他们能够不时地发出会心的微笑。

任何一个项目,无论是一篇论文的写作,还是一本书的编写,都离不开他人的支持和帮助。在此要特别感谢多年来与我们一起工作、一起分享过困苦和胜利喜悦的同学们。此外,很多同事也提供了他们的学生对于本书第 1 版内容是否能给他们提供帮助的反馈信息。美国心理学协会(American Psychological Association, APA)的

工作人员们(Julia Frank-McNeil, Mary Lynn Skutley, Peggy Schlegel, Susan Reynolds, Susan Herman)对本书一些冗长的、不清晰的陈述进行了修改,在此表示感谢!最后,还要感谢一直以来默默地给予我们支持和帮助的家人:Jan Cone 和 Tom Barton,在本书第 1 版的编写和第 2 版的修订过程中,他们在承担了各自本职工作的同时,分担了更多的家庭事务,使得我们有更多的时间来完成本书。

序　言

致研究生的公开信

祝贺大家进入研究生院!

人人都有不同的过往,但大家现在都开始了研究生阶段的学习。攻读研究生学位多半会赢得掌声,虽然偶尔也会充满艰辛。生活可能变得拮据困难,但有同学们风雨同舟,便也是苦中作乐。不久之后,大家将会遇到一生的知心伙伴。在这个优秀的团队中,大家也将学会如何面对世间之磨难困苦。总之,祝贺大家!

谈到选题和学位论文的撰写,我的建议可能并非完全适用,毕竟我已经毕业半个多世纪了。但从另一个角度讲,在这个领域多年工作的经验,让我能够站在巨人的肩膀上看待问题。并且,我愿意将自己所领悟到的心得拿出来和大家一起分享。

建议一:科研机构推动着科学和专业的发展,可以将读研这段时间看做是驶向目标的准备和学习阶段。完成学位论文,可能是大家对科学做出的最重要贡献,它也许是你到目前为止做过的最大的研究项目,但时间会证明其价值所在。

下面具体谈谈对学位论文写作的建议。

建议二:学位论文的选题应该与自己的智力风格相契合。社会科学研究包含从质化到量化的精确度范围,心理学家认为每个人都有所谓的"**暧昧容忍度**"(tolerance for ambiguity),因此要考虑认知匹配的因素。愿大家选择自己能够驾驭的、严谨明确又令人雀跃的研究主题。

建议三:学位论文需要构建和完善已经成型的研究领域。选择特立独行的主题可能是创造性的需要,但这并非明智之举! 因为这样你会失去与其他从事相关研究的同行交流的机会,同时,能得到的指导也会微乎其微。导师也并非无所不知,这样"特立独行"的选题对学科发展也是意义甚小。相反,这只会使学科缺少连贯性和整体性。

建议四:你与指导者之间是一种合作关系。当你决定选某个指导者时,和他取得联系,并请教他的研究想法。了解到指导者的研究想法后,你应该先独自考虑其中的一两个,然后把你的研究方案告诉他。如果一切顺利,他可能会说:"我一直期待你对该领域兴致盎然,这个课题是很好的开端。送给你一些我关于该领域的研究论文。是不是该增加控制组呢? 增加那个额外条件合适吗? 期待我们今后进一步的合作,我确信我们彼此都能有所收获。"如果你发现自己无法通过上述方法与预期的指导者建立关系,那就再去寻找其他指导者吧。

建议五:在撰写论文的时候,要在允许的范围内尽可能简洁。学校可能会对论文的最低字数有约定俗成的要求,但是几乎可以肯定,这一标准超出了你撰写时所需要的简洁程度。当然你不需太过介意,一个我指导过的最优秀的学生,他的论文没有超过20页。

接纳导师针对文章语言表达的具体细节提出的建议。我写论文时参考 Kenneth

Spence 学到的写作知识比我在大学的任何英语课程学到的都要多。

建议六:学到的研究方法远比研究本身重要。社会科学的多数“真理”时隔五年影响力就会减半,而研究设计与方法具有更强的生命力。可以尽量避免使用**操作性定义**(operational definition)这样的敏感词,但无论如何,都应该解释清楚诸如“自我概念”“记忆表征信息”和“分离焦虑”这样的概念,以保证你提供的信息足够使他人能重复你做过的研究。

对那些偶然的惊奇发现持质疑态度,它们可能只是由于抽样偏差(Ⅰ类错误)所致。多年以来,很多由意料之外的结果建构起来的理论,都随着理论的丰富及观点的不断扩展而最终土崩瓦解。因此,出人意料的结果往往需要引起额外关注,除非该结果得到重复证实。

下面与大家分享三个总结性的建议。

建议七:督促自己成为科研团队中的一分子。积极参加科研座谈会——尽管多数课题与你的研究领域无关,但思考其他领域的问题,会让大家对自己的研究更加明晰。通过与报告人的交流,能够多了解其研究的发展现状和该领域的未来走向。此外,要把握机会去参加学术会议,多与你拜读过作品的学者接触,了解该领域的国际组织并加入其中,成为学生会员。

建议八:逐步专业化。未来的 4 ~ 5 年间,大家将浸泡在实验室和图书馆中,但毕业后还得谋生。要对此做好准备,了解即将进入领域的人情世故和商业技巧。

尝试在学术会议中提交论文和演讲,或许有机会接触到你未来的同事甚至老板。你也会得到关于该领域研究和应用最前沿的信息。

学会准备和撰写投稿论文。

建议申请研究奖学金。如若有幸获得,你会得到实实在在的收益,研究生涯也会如鱼得水。而且,申请奖学金的过程也有利于培养今后申请基金项目的能力。

建议九:这是我极不情愿说的建议,但事实如此。研究生课程不及格怎么办?多门功课不及格从而延迟毕业怎么办?被强迫去做你不喜欢的科研(或治疗、教育)怎么办?好在一切自有神庇佑,如若你陷入困境,请自我审视并思考下面两个问题。

首先,你是否真的想从事这一领域的研究工作?如果答案是肯定的,那第一步就是找到一件让你欢愉无比、不求回报的事情;第二步才是得到报酬。也许你现在的情况并不是这样。

其次,自己是否能够胜任这项研究工作?事实上,人的一生发展的过程就是不断完善和自我评价的过程。如果通过正确的自我评估,发现自己在此领域很难辉煌,那么不妨去寻找人尽其才之所。

最后,建议十:本书提供了许多有益的信息,详细地阐述了上述建议。大家会从中看到许多感同身受的观点和看法。本书还对许多大家未曾(或不敢)思考的问题给出了解答。希望大家能够从中获益。最理想的是,如若你有幸成为教授,这些建议也会给你的教学带来便利。

再次祝贺你们!愿你们在研究生院度过短暂而又美好的时光!

Gregory A. Kimble
杜克大学名誉教授

目录

1

什么是学位论文？为何要编写本书？

写作这本书第1版的目的就是为了帮助心理学以及相邻领域的研究生能够更加顺利地完成学位论文写作的全过程。同样,我们也想把这本书作为自己指导学位论文写作50年心得的一份档案资源。我们都传授了很多有用的知识(当然有一些也不是特别有用)给研究生们,以帮助他们完成论文写作过程。在我们系统地整理这些知识之前,我们和其他导师一样,在不停地对每一个在学位论文写作上毫无经验的研究生重复相关知识的传授。本书的第1版很好地体现了如何把论文写作过程变得简单和有趣的教学理念。在这一版中我们加入了一些上次忽略的和心理学领域中最新出现的东西。尽管本书的主要目的是给高年级的研究生提供关于从事心理学和相关领域学位论文写作的建议,但其中许多内容也同样适用于其他领域的新人。

本书是如何组织的?

本书的14个章节提供了写好学位论文应该具备的许多细节知识。当然这些章节也不可能面面俱到。本书将充分满足对研究生教育的需求。我们在书中并不解释任何关于研究设计和统计方面的具体内容,取而代之的是帮助学生们应用那些他们已经学过的知识,来尝试着进行研究和论文写作。

在本书的第1章,我们简要地介绍了学位论文究竟是什么?学位论文是什么样的?为什么需要完成学位论文?第2章,介绍了如何评估研究生的准备情况,还包括如何获知学位论文写作的一般标准,如何协调学位论文写作过程中研究生的生活等许多建议。第3章,帮助研究生评估完成论文写作需要的时间,预期并知晓如何处理大量的可能不利于论文写作的事情。第4章,探讨如何确定一个研究方向,并将其完善成一个可研究的问题和假设,因为这些是与导师合作时最常进行的工作。同时,第4章也包含了如何选择答辩委员会主席及其成员(与国内的答辩机制不同,美国的答辩委员会主席和其中的成员需要申请进行答辩的学生自己来组织——译者注),以及学位论文的一般常规性的建议。第5章,提供了一个对学位论文要求的总揽。第6章,文献回顾。第7章,讨论了研究技术、研究的伦理学和知情权等问题。第8—10章,涵盖了如何选择合适的测量,选择统计处理方法,如何进行数据搜集和数据分析,等等。第11章,如何在论文中呈现研究结果?第12章,结果的讨论部分。第13章,开题和答辩的策略。最后,第14章,讨论如何使你的论文符合要求,并提交到答辩委员会,以及论文的最终发表。

在大部分章节的最后,我们提供了补充资源,以便大家更多和更深入地了解这一章的主题。尽管我们的网站不是永久的,但是我们也提供了一些其他有用的而且会长期存在的网站。当我们推荐的那些补充资源的主题或推荐的理由不是很清楚的时候,我们给出了补充资源的分类主题,或者对某些资源加注了简要的介绍,以供读者参考。

部分章节附有一份清单,可以把我们的建议转化为具体步骤,帮助你继续进行研究。你可以修改这些清单的内容,以使其适合你的研究需要,当然也可以用它们来跟踪你的工作进程。最后,我们在附录中加入了美国心理学协会(APA)制定的道德准则。

本书的主体内容涵盖了学生们在撰写学位论文过程中可能遇到的一系列问题。相信你能够按照书中的章节把这些建议应用于论文写作的每个段落之中。或者你也可以在开始写作论文前浏览全书,这样会使你对本书有个大概的了解,然后找到自己需要的部分,而不是按照我们编写的顺序来进行阅读。

现在让我们回到本章标题里的第一个问题吧:学位论文究竟是什么?首先,我们来解释一下学位论文的定义;然后再讲讲学位论文的历史。最后,来看一看心理学学位论文大致是什么样子的。

定义,区别和功能

在基础英文词典里,对硕士学位论文(thesis)或博士学位论文(dissertation)的定

义表明，它们一般没有什么区别。韦氏在线词典（Merriam Webster Online）对博士学位论文的定义为“对一门学科的拓展性的书面写作；**具体来说就是**：为获得博士学位而提交的文章。”对硕士学位论文的定义为“一篇包括独创研究结果以及证实了一个特定观点的文章；**特别是**：一个申请者为得到一个学位而写作的文章。”类似的，AskOxford. com，牛津英语词典网站把博士学位论文定义为“为获得大学学位或文凭而写就的长篇文章”，硕士学位论文是“作为获得大学学位的一部分，并包含个人研究的学位论文或长篇文章”。可以看出，二者在以上的解释中没有明显的区别。

而在美国的大学里，通常把为了获得硕士学位而提交的论文称为 theses，而把申请获得博士学位的论文称为 dissertations。但这并不是规定的叫法，有些教师把博士学位论文也叫做 thesis。本书通篇使用的“博士学位论文”是指一个独创的、为获得博士学位（EdD、PhD 或 PsyD）而撰写的心理学及相关领域的实验研究报告；“硕士学位论文”指为获得硕士学位而撰写的实验研究报告。

尽管术语的解释是那么模糊，但是绝大多数教师都认可它们的一般职能。撰写一篇学位论文的一个主要目标是展示学生在一个重要课题上进行的独立研究，并做出自己的贡献的能力；另一个目标是评估学生在特定领域的研究水平。有些人认为学位论文的写作是个人驾驭研究技术的体现，包括研究设计、统计学等理论的应用能力。因此，可以看出学位论文的写作过程是对自己作为研究者是否能够独立驾驭一项研究的检验。而且，博士学位论文也应为增加相关领域的新知识做出卓越的贡献。当然，硕士学位论文也应做出很多贡献。硕士学位论文不过多地要求原创性，然而提交者要对论文的内容给予更多的说明和检验。例如，一些硕士学位论文提交者会重复一个前人已经做过的系统研究。另外，硕士学位论文的提交者也很少像博士学位论文提交者那样完全独立地进行研究操作。

此外，博士学位论文和硕士学位论文有着相同的训练基础。要进行学位论文的撰写，就要在自己的研究领域中学习和完善研究技能和相关知识。也要扩展你的批判性思维和综合能力等各方面的能力，并且能够与他人清晰地进行专业交流。完成学位论文的目的不仅仅是为了毕业，其实最理想的结果是加深对学科的理解，并让你对实证研究充满乐趣，甚至为自己在完成学位论文过程中所取得的成就而感到自豪。

学位论文看起来是什么样的？

大家可能已经开始思考关于学位论文的许多问题了，如研究必须使用特定的方法吗？论文的长度应该是多少？应该遵循什么样的格式？对写作风格有什么要求吗？现在就让我们来回答这些问题。

首先，关于研究方法。尽管本书假定你的研究将包含定量的研究，但是并不是所有的学位论文都必须如此。心理学中传统的最高学位：博士学位（PhD），实际上总是要求采用实验研究。一些应用心理学提供另一种最高学位：心理学博士学位（PsyD）。心理学博士学位的教育计划是为了培养从业医生，多数是临床心理学方面的心理医生。但是，这并不能排除也要求申请者写作出具有独创性贡献的学位论文的可能。当然，许多医学导向的学校允许使用非实证性的学位论文来替代。桑奇

斯-胡克莱丝和卡什(Sanchez-Hucles and Cash,1992)调查了 40 名“韦尔模型”(Vail-model)教育项目(包含了美国绝大多数的临床心理学的博士学位教育项目)的从业导师,75%的受访者表示他们接受非实证性的学位论文,如理论型的文章、项目评估和临床案例报告等。但有趣的是,这个方向几乎 50%的学生会选择进行实证性研究,以确保他们能获得相应的学位。

从论文长度上来讲,不同的大学有不同的规定。尽管我们在这方面对硕士学位论文没有确实的数据,但是我们确实知道博士学位论文的一般长度是多少。最近,我们从电子论文库(Digital Dissertations database)中随机抽取了 100 份 2001—2002 年的博士学位论文。心理学学位论文(博士论文)的长度平均在 190 页(不包括附录超过 450 页的论文),最少的 89 页,最多的 612 页①,抽样的标准差是 89.6。如今 85%的心理学博士学位论文的长度从 63 ~317 页不等,其中包括正文、表格、附录以及相关的其他部分内容。我们的经验是硕士论文一般有博士论文的 2/3 那么长就合适了,也就是 127 页左右。

学位论文有其固定的格式,每个学校都会有严格的规定。我们建议你与负责此事的相关人员联系,并从他们那里获知相应的规定和要求。他们可以给你提供一个关于对学位论文格式进行详细说明的书面材料或是网址。另外,他们也可能给你提供其他一些有用的信息,例如论文提交的日期、要求、装订费用等。Publication Manual of the American Psycholouical Association(APA,2010)②对心理学学位论文的一般格式做出了相应的要求。本书就是按照 APA 格式编写的,而且许多行为科学的期刊也遵循 APA 格式。其实,全书自始至终都广泛地参考了 APA 格式的要求。

尽管 APA 格式是一个被广泛使用的格式,但也不是在所有的心理学院系都适用。甚至,就是在那些接受 APA 格式的地方,也经常存在一些差异,例如,在引用和表格的布置中的差异等。大家在论文写作初期就了解这些差异会比较好,我们建议大家应先了解你所属学校的规定和要求。从开始动笔写作时就按照规定的格式进行,这样可以帮你节省大量的时间和精力。

了解学位论文应该遵循什么格式的最好方法是查阅一下导师指导过的其他学生的学位论文,这可以给你提供一个最好的参考样本。也可以请教导师都有哪些建议,导师可能会让你参考与你所做研究最相关的学位论文,但也要知道那些学位论文的质量也是有高有低。本书的第 4 章明确地介绍了如何选择答辩委员会主席来主持你的学位论文答辩会,所以此时你不必为这个问题担心。

那么,在论文的最终稿里应该包含哪些内容呢?绝大多数的博士和硕士学位论文与期刊论文很相似,只是学位论文应该包含更多的细节,应该包括目录、文献回顾、研究方法、结果、讨论、参考文献和附录等部分。参考文献应该包含所有在正文中引用的文献,附录要提供研究需要的补充材料,如器材图、实验知情同意书、数据记录表等,有时甚至还需提供研究的原始数据。

---

① 感谢琳达·伊萨克收集和分析这些数据。

② 本书中文版《APA 格式——国际社会科学学术写作手册》已由重庆大学出版社引进出版。——编者注

为什么要把学位论文摆在首要位置？

有很多理由需要学生完成一篇学位论文，但是归根到底是为了获得相应的学位。当然，可能还有许多其他的原因，如通过了学位论文并获得了学位，就可能找到更高薪水的工作，或者成为自己专业生涯的开端。其他的一些原因，诸如：论文的撰写工作为思考和解答概念上、方法学上和实践中出现的问题提供了许多的机遇和挑战；进行相应的研究也是一条更好地激发自己好奇心来发现新的心理学或行为学现象的途径；完成一篇学位论文还可以使你因为承担一项复杂的具有挑战性的任务而得到自我满足，等等。

目前，你已经处于要开始完成一篇学位论文的位置上了。不妨认真想一想自己要如何看待完成学位论文这项工作：它仅仅是你在获得学位的道路上必须完成的一件事情吗？你在寻找最容易和最快捷的途径吗？你是不是渴望着在这样一个研究过程中积累一些有益的经验？你觉得它会和自己毕业后计划从事的工作有关吗？你有机会继续自己之前的研究并且对新的问题进行解答吗？

从认知-行为治疗学家那里可以知道，对自己的研究方式进行反复的考虑，有助于从已经学到的东西中获得经验并增加乐趣。如果你对心理学的一些方向或心理学研究的过程充满了好奇心，那么我们真为你感到高兴！你可能会对自己的研究自始至终展示出充分的耐力，甚至在这一过程中还能发现不少的乐趣。这对注意力集中在研究工作上有着积极的作用，并且会把遇到的困难看做挑战，而不是障碍。如果你不喜欢学习的话，可能你现在不会理解这些。进行学位论文的撰写工作，在很多情况下，可能是你一生中最后的学习经历。

找出对你有效的部分

在结束本章之前，给大家一个忠告：我们提供的一些建议可能并不适合所有人。不是所有的策略对每一个学生都适合，也不是每一个导师或学院都使用相同的学位论文指导方式和评分标准。不同的研究生有许多不同的论文完成方式，关键是找到对自己有效的方法。另外，因为每个地方的规则和传统也都不同，对书中的内容要报以怀疑的态度，然后看看书中提供的信息与你所在学校的要求是否一致。记住，完成一篇学位论文，从某些方面来说，就要遵循一定的规矩和要求。与大多数的规矩一样，其中的一些可能是专断的和缺乏根据的，但如果你预先知道了这些规矩和要求，而不管它是否有意义都按照这些要求去做，你将会轻松地完成论文写作过程，并且从中获得很多的乐趣。

从另一方面来说，也不要被学位论文的完成过程吓倒。是的，学位论文的写作会是一个很长的过程；你需要投入很多的精力；之前你从来没有接触过任何类似的东西；有时这个过程可能很恐怖。但是，我相信你和其他许多人一样，能够完成学位论文并最终获得学位。另外，导师和答辩委员会成员也将在这个过程中帮助你。

我们的一个学生说得好："当要结束研究生生涯并且完成学位论文时，一个想

法在脑海中挥之不去,即我取得的任何一个成就都由两个要素组成:①把一件事情分解为几步来做;②家人、朋友和团体的支持"(Dionne,1992,p. iii)。

我们知道你已经充满热情,不然也不会阅读这本书。你已经准备好要付诸行动了吗?那么,让我们转入第2章看看你都做好了哪些准备,并且你还能再对哪些细节准备得更加充分。

2

开始：评估工作准备

在第1章中我们建议大家重新审视对即将开始的学位论文写作所持的态度。在这一重大写作项目开始的时候和进行的过程中，始终保持一个正确的思维习惯是十分重要的。准备工作其实也有很多种，本章的任务就是帮你评估它们。我们的目标是帮助你正确评价那些能够让你顺利地完成工作的技术。

你准备好了吗?

为了让你能够马上开始进行自我评估,我们准备了一项测验。到目前为止,你已经知道了为什么要完成毕业论文,而且也说服了自己努力地去完成这件事。尽管如此,你真的准备好要开始工作了吗?完成下面的研究准备清单(见专栏2.1),可以很好地帮你回答这个问题。完成下面的测验的时候一定要确保真实地回答所有问题。记住,这个测验只是为了测试你达到了哪种准备状态,一定要诚实地做答。一个否定的答案能够为你提供很多有用的参考内容。将答案写在一张纸上,然后我们在下一个部分来具体地解释。

专栏2.1

研究准备清单

是 否 你的写作水平如何?

☐ ☐ 1. 你是否从教授那里得到反馈说你的文章很容易理解并且逻辑很清晰?
☐ ☐ 2. 你是否正确的保持语法一致性?
☐ ☐ 3. 你是否能保持正确的拼写?
☐ ☐ 4. 你是否很熟悉APA格式,在写作过程中只需要偶尔查阅?
☐ ☐ 5. 你知道如何使用一个文书软件吗(例如,WordPerfect,Word)?
☐ ☐ 6. 你能有效地组织自己的论文吗?
☐ ☐ 7. 你在开始写作之前会准备提纲吗?

是 否 你掌握了必要的研究方法吗?

☐ ☐ 8. 在过去的3年中你是否已经学习过至少两门研究生统计课程?
☐ ☐ 9. 你是否学过研究生测验编制或测量理论课?
☐ ☐ 10. 你是否修过研究设计课程?
☐ ☐ 11. 你在研究生院的时候进行过文献阅读训练吗?
☐ ☐ 12. 你在研究生期间参与过实证性研究吗?
☐ ☐ 13. 你是否至少会使用一种统计软件?

是 否 你具备必要的研究准备条件吗?

☐ ☐ 14. 你是否至少同3个人以上探讨过关于学位论文写作经验的问题?
☐ ☐ 15. 你是否阅读过其他同学已经完成的学位论文?
☐ ☐ 16. 你每周是否最少能花10~20个小时在学位论文写作上?
☐ ☐ 17. 在接下来的12~18个月里是否有充裕的时间?

是 否 你掌握论文写作的相应技巧吗?

☐ ☐ 18. 到目前为止,你有一个可以不受打扰的写作空间吗?

专栏 2.1 (续)

☐	☐	19. 你有足够的参考资料(图书馆资料或数据库)吗?
☐	☐	20. 有教师或导师对你进行基本指导吗?
☐	☐	21. 你能获得家人或配偶的支持吗?
☐	☐	22. 你自己拥有或者可以使用电脑吗?
☐	☐	23. 你的键盘输入技巧很熟练吗?
☐	☐	24. 你知道如何使用所在大学图书馆的数据库、文献检索和仪器设备资源吗?
☐	☐	25. 你是否掌握行之有效的时间管理技巧?
☐	☐	26. 你是否可以有效地和他人沟通交流?
☐	☐	27. 你是否知道所在学校学位论文写作的正式要求?
☐	☐	28. 你是否知道学位论文写作过程中的一些约定俗成的规定?
☐	☐	29. 你有没有问过他人进行研究的开销是多少?
☐	☐	30. 你是否有足够的支持完成学位论文写作的经济来源?

这份清单不是一个实证性的测量,所以不要以为它提供的答案可以完全代表你对进行一个复杂项目而作出的准备的程度。相反,要把它作为一种自己是否能够轻松地完成学位论文写作这一过程的参考。这个表格实际上是我们对你是否具有完成学位论文的技能和资源以及时间分配等情况的分析。同时也给你提出了一个警告:不要让一个“否”的答案成为引发焦虑的刺激物。一个“否”并不代表着你就不能完成学位论文。例如,我们的一个学生,她在上表中至少得出了 4 个“否”的答案,但是仍然成功地完成学位论文。而且也不是全部得到肯定的答案就意味着你能在学位论文写作的过程中一帆风顺。确切地说,用“否”这个答案来提醒你预先进行充分的计划和准备可以帮助你避免在写作学位论文的过程中出现一些潜在的问题。下面我们来讲讲这些准备工作都应该包含哪些内容。

对测验结果的解释

在分析你的准备程度之前,先让我们来看看与上述清单联系密切的一些东西。

写作 写作和研究方法可能是确保你能顺利地完成一篇学位论文最重要的技能。如果你对清单中的题目 1 和 2 给出了否定答复,那就真的需要在开始论文写作之前认真地考虑一下补救的方法。语言组织能力表明了你思维的清晰程度。如果你不能组织好语言(题目 1),那么你可能在文章的逻辑结构上存在不少麻烦。不幸的是,没有简单而又快速地训练这些技能的方法。所以如果你觉得自己不能很好地组织语言材料,一定要停下来重新考虑一下。也许你可以从相应的语言组织类课程或者培训上受益,去文学院查询一下。你上过逻辑学的课程吗? 如果没有的话去哲学院找一找。语法使用的正确与否(题目 2)要比语言组织技能更加具体有效。如果你在校期间一直在句子结构、段落组织、写作技法和词汇选择上有困难,那么一定要在开始写作学位论文前认真的想办法补救。尽管一些学校允许对学位论文进行后续的编辑,但是绝大多数学校都不允许论文中出现严重的语法错误。除非你的母

语不是英语,不然教师们还是希望你能自己动笔写作论文,而不是请人代笔。

你可以运用一些方法来提高自己的语法技能。首先,需要学习关于句法、语法规则和标点的知识。有了这些句法和标点的基础知识,下一步就是从已经写成的文章中选出一部分让语文老师帮你检查一下。Word™软件包含了很多语法和拼写检查选项,应该可以满足你的需求。另一个独立版软件 StyleWriter™也非常好用,你很容易在因特网上找到它。

这些软件可以指出你的文章是不是违反了语法规则,犯了标点符号的错误,运用了太多的复杂句、否定句或者被动语态等问题。使用这些程序几乎就和雇佣了一个专业的编辑来审阅你文章的效果一样。你会为这些检查工具对文章分析得如此透彻而感到震撼。即使你自认写作水平不错,也会从中受益颇多。我们人人都有一些不好的写作习惯,只是我们都不自知罢了。

举一个语法检测程序的例子,专栏 2.2 是一个简单的草稿写作样例。

专栏 2.2

StyleWriter™的段落修改建议

Writing. Writing and methodology[1] skills are probably[2] the most important skills you need to complete a thesis or dissertation. We'll discuss methodology in a moment. First, let's talk about writing. If you said "no" to Items 1 and 2, you should seriously consider additional[3] preparation before undertaking your project. Organization skills in writing (Item 1) are an indication[4] of how clearly you think. If you do not organize well, you might have trouble thinking in a logical fashion. And, unfortunately, there are no quick and easy programs to teach this skill. So, if you have reason to believe you do not organize verbal[5] material well, stop and reconsider. You might benefit from specific coursework or tutoring in organizing and sequencing written material. Examine offerings in the English department of your school. Have you ever taken a course in logic? If not, look for one in the philosophy department. Grammatical prowess[6] (Item 2) is somewhat[7] more specific than organization. If you have consistently received input[8] during your university schooling that you have problems with sentence structure[9], paragraph organization, and the proper choice of words, give some serious thought to remediating before beginning your thesis or dissertations. Although some schools permit[10] editorial consultation on theses and dissertations, most stop short of allowing the kind of[11] input[12] serious grammatical deficiencies[13] would require[14]. Unless English is not your native language, faculty will expect you to write the document yourself, not have an editor do it.

Words = 241　Style Index: 53 Average　Ave. Sent: 14 Excellent　Passive: 0 Excellent

注意:使用这个软件的时候,你可以点击带有阴影部分的词汇看到 StyleWriter™给出的解释,StyleWriter™建议使用其他词汇更替它们。以前面的 9 个带阴影的词汇为例解释如下。

1. 复杂词

一般用简单词替代

尽量使用简单和熟悉的词。简单词让你的文章更有可读性。使用简单词汇可以让你的意思表达得更清楚,这样读者的注意力可以集中在信息上。

专栏 2.2 (续)

例如:Please endeavor to ascertain the truth.

Please try to find out the truth.

2. 重复性

避免华丽的修饰语

通过简化繁琐的段落来提高写作能力。可以通过用一个不失原意的简单句代替一段表达。将句子中不必要的词去掉可以使其更加精炼和清晰。节约用词是一个好文章的标志。StyleWirter™把繁琐的表达通过阴影的形式标出,然后提供一个简短的可供替换的材料。例如:

繁琐的	建议的
at a later date	later (be specific)
due to the fact that	because, since, for, as
in order to	to
in relation to	about, in, with, towards, to, on (or omit)
prior to	before

3. 复杂词汇

建议:added, extra, more, another

4. 复杂词汇

建议:sign, signal, mark, hint, suggestion

5. 混淆词

建议:aural = of the ear or hearing

oral = of the mouth, voice or speaking, spoken

verbal = of words or speech, usually spoken

6. 误用词

建议:prowess = valor or daring—does not mean skill or competence

7. 重写/覆盖

建议:一般情况下可以忽略“somewhat”。如, somewhat older = order

8. 术语/摘要

建议:不要过度使用[具体化或编辑掉]

9. 术语/摘要

建议:不要过度使用[具体化或编辑掉]

注意:在 StyleWirter™中 Ave. sent = 平均句子长度。此软件根据所选写作类型评估句子长度,数值越大,代表句子越长,一般认为也越难理解。

从中可以看到,StyleWirter™为这个样例提供了详细的分析。你的写作水平可以通过 StyleWriter™的分析来获得提高,如专栏 2.3 展示的那样,注意专栏下面的字数统计。一般 StyleWriter™的修改建议会比这个例子更少一点。你不会从答辩委员会和导师那里得到这么丰富的修改反馈,那些“大忙人”给出的建议只能是你的写作水平需要提高。他们可能会指出一些特定的固定词组搭配、动词-主语不一致和介词错误,不过要知道,他们的主要工作是教授研究技能而不是写作技法。当然也不是所有人都能从文字软件程序中获得有用的反馈。而且即便在时间允许的情况

下,也不可能有人把上千条语法规则记在脑子里并在需要的时候进行提取和使用。罗素·萧(Shaw,2002)写过一篇有趣的文章阐述了语法检查软件的正反两个方面,你可以在网上找到这篇文章。

切记,本书的写作并没有使用正式的写作方式。通篇充满了像“你的”这样的词汇和缩略,例如使用“and so on”等。这些都是不可以在正式写作中出现的,也不可以在学位论文中出现。请不要使用本书的写作风格作为你的写作参考。

对于提高写作水平的最后一个建议是找一份在你的研究领域中最有权威的期刊作参考。这份期刊应该发表了很多与你要做的实证研究相接近的文章。从最近的期刊中选择几篇文章,弄明白其中一个的大致结构。看看作者是如何使用标题和副标题来组织全文的。

在本节中,请注意不同的段落是如何进行承接的,信息呈现的逻辑是怎样的。最后,注意句子的结构,确保自己足够了解这些细节。提高写作水平的最佳方式是准备自己的提纲,然后检查其组织和结构。在提纲下面留出足够的空间并把每个标题下的文字都复制过来。要逐字逐句复制过来,包括那些支持或扩展论点的句子也要一起复制过来。之后你将对自己的写作水平有一个更深刻的认识。可以对导师给你做参考的、那些特别出色的学位论文进行这样的结构分析。

拼写(题目3)和写作风格(题目4)是更加具体的语法技能。尽管它们很重要,但是幸好即使拼写水平很差也不会对行为科学家造成严重的影响。因为如果你拥有一台电脑,你可以把自己的文章输入进去,使用拼写检查程序挑出其中大部分的错误,没有电脑的话可以雇人来帮你校对,美国心理协学会也是这样要求的。一些计算机程序(如 StyleEase for APA Style™)可以完成大部分的校对工作。如果连这样的程序也没有的话,就雇一个熟悉 APA 格式的人来帮你编辑。最好买一本 Publication Manual of the American Psychological Association(APA,2010),然后自己学习掌握这些格式。任何一个有能力完成毕业论文的人都可以通过学习这本手册掌握 APA 格式。另外 APA 还出版了一本学生指南也很有用,尽管书中的主要内容是关于期末论文的写作而不是关于学位论文写作的。

专栏 2.3

StyleWriter™修改过的一段

Writing. Writing and methodology skills are the most important skills you need to complete a thesis or dissertation. We'll discuss methodology in a moment. First, let's talk about writing. If you said "no" to Item 1 and 2, you should seriously consider more preparation before undertaking your project. Organization skills in writing (Item 1) are a sign of how clearly you think. If you do not organize well, you might have trouble thinking in a logical fashion. And, unfortunately, there are no quick and easy programs to teach this skill. So, if you have reason to believe you do not organize verbal material well, stop and reconsider. You might benefit from specific coursework or tutoring in organizing and sequencing written material. Examine offerings in the English department of your school. Have you ever taken a course in logic? If not, look for one in the philosophy department. Grammatical skill (Item 2) is more specific than organization. If you have consistently received feedback during your university schooling that you have problems with sentence structure, paragraph

专栏 2.3 (续)

organization, and the proper choice of words, give some serious thought to remediating before beginning your thesis or dissertation. Although some schools allow editorial consultation on theses and dissertations, most stop short of allowing the help serious grammatical deficiencies need. Unless English is not your native language, faculty will expect you to write the document yourself, not have an editor do it.

Word = 236 Style Index: 21 Good Ave. Sent: 14 Excellent Passive: 0 Excellent

注意:在 StyleWirter™ 中 Ave. sent = 平均句子长度。此软件根据所选写作类型评估句子长度,数值越大,代表句子越长,一般认为也越难理解。

如果你所在的学校不使用 APA 的论文格式,那么一定会有人给你一份格式要求表、说明书或者说明网站来告知你相应的格式,你可以去图书馆查找或者问问教务秘书。在因特网上快速地搜索一下,你还可以找到与 StyleWriter™ 相类似的文字软件供你使用。

方法学

题目 8—13 涉及你在方法学上的准备是否充分。在大多数院系里,一般来说选修 2 门或 2 门以上的统计课(题目 8)就足够很好地完成学位论文了(当然,如果在应用心理学研究的分析过程中涉及用图表表示结果时,图表的内容又大量地受到语言描述的限制,情况就复杂得多)。如果你连续在两门统计课程上都没能很好地掌握统计理论和灵活地使用统计工具,那么你可能需要进行一些辅助的学习。需要掌握的统计工具等内容见本书第 9 章。看一看,你熟悉它们吗?是不是只需要快速地复习一下?

如果离你修完统计课程有一段时间了,那你可能想系统的评估一下已经掌握的统计技术。假如研究生课程的综合考试里面也有统计考试的话,问问导师你能不能参加现在的考试。向他/她解释一下你想通过考试来了解一下自己是否需要重新上一些补充课程来巩固以前学过的知识。

在患上学位论文写作过程中的"统计学恐惧症"之前,记住我们所说的是"熟练"掌握统计方法和工具。我们的意思不是指你得在统计课上拿到 A + 这个成绩,或者你能很容易地得到需要的统计材料。在我们的教学经历中发现,绝大多数学生都会觉得统计学是最令人生畏和最具有挑战性的部分。即使是那些拿到 A 的学生也会发现他们需要复习一下过去的学习材料,或者经常也会为该使用哪种统计方法而头痛。不过,只要有足够的指导,他们完全可以处理好统计工作。只要你已经掌握了基础统计理论,那么只需要足够的指导就可以了,但是如果没有掌握这些统计学知识,那就需要接受额外的训练或培训。数以千计的心理学家通过相应的统计训练克服了他们对数字的恐惧,你也可以做到。

如果你计划使用自己编制的量表,那么对题目 9 的肯定回答就显得非常重要。你要研究的领域没有合适的测量工具吗?例如,你计划要调查临时工的态度吗?是否对男同性恋的约会关系感兴趣,结果却发现没有合适的工具呢?如果对这些问题的答案是肯定的,那么你需要熟练地掌握测量开发技术或者测验结构编制的理论才行。量表的结构和效度是纷繁复杂的,而且也不能够轻率地使用。即使你的研究使

用的是前人编制好的量表，最好也要修订一下原来的量表，然后再加以应用。想编制出适合自己研究的好量表，不仅需要熟悉心理测量理论，还需要掌握其他很多相关方面的知识。如果对自己在这方面的准备情况还拿不准，请参见第 8 章内容。

你修过研究设计课吗（题目 10）？这对使用何种研究方法来提高研究的内部效度和外部效度非常重要。如果你对这些一窍不通，我们强烈建议你复习一下研究设计的理论知识。

研究设计课的另一个目标是让你能够使用研究性语言正确地表达自己的观点。学生们在表述**方法问题**时经常会出现**混淆**，其还有关于研究设计的信度的问题，或者混淆了**因变量**和**自变量**。正确地使用专业术语，在论文中讨论你和他人的研究方法是非常重要的。另外，如果不能在答辩中正确地使用这些术语，则意味着你向答辩委员会成员们发出了一个根本不知道自己在说什么的信号。这些有时会引起不必要的麻烦，而且有时还会对提问造成相反的效果。幸运的是，如果你确信自己已经正确地理解和使用研究术语的话，上面所有的问题都可以避免。

题目 11 是关于阅读和评估研究文献的经验。不是所有的已发表文章的研究方法都很严格。要进行一项好的研究，你的工作之一就是去粗存精（我们当然希望不好的研究论文从来就不曾发表），然后确保你的研究是那些"好的"部分。此外，你应该对别人（在文献回顾部分）和自己的工作（在讨论部分）报以批判的态度。大多数事情都是熟能生巧的，因此在研究生课程中对实证文献的批判性阅读和评价可以提高思维的敏锐性。梅尔佐夫（Meltzoff，1998）总结了很多存在内部问题的研究，并附上对这些问题的解答，如果你需要更多的练习可以看看这篇文章。

从一定程度上说，对题目 12（研究经验）的肯定回答可能涉及你对之前题目 8—11 这 4 个问题的重视程度。特别是如果你已经为今后的研究工作进行了一些基础研究的时候，就会对研究的方法学问题比较熟悉了。如果你没有那么多经验，或者计划从事另一领域的研究，可能就不会准备得特别充分。这样的话，就要对自己需要使用的统计学方法和测量学理论进行一个全面的评估。

绝大多数的教学课程都让学生将处理数据的苦差事交给计算机软件。其中最受欢迎的统计软件是 SPSS，SAS 和 Systat 等，看看哪些是你能使用的。如果你现在还不知道需要在研究中使用哪种测试或实验程序，那就尽量熟悉所有可用的资源。在完成研究设计之后，就会知道自己需要分析哪些数据、使用哪种工具了。

总之，在学位论文的整个过程当中，写作和方法学技术是成功与否的决定性因素。在继续下面的工作之前，让我们看看这份测验清单上其他题目都有哪些含义。以题目 16 和 17 为例，一些学位申请人无法理解每周至少投入 10 ~ 20 小时的工作时间，而如此投入还要持续 1 ~ 2 年更令他们目瞪口呆。事实上，这个时间对于完成一篇博士学位论文来说还只是保守的估计。戴维斯和帕克（Davis & Parker，1979）发现 80% 的博士学位论文需要每月投入 175 个小时的工作时间，持续 11 ~ 19 个月才能完成。当然这个时间也不是一成不变的，题目 14 和 15 就是为此而存在的。在你的学校里，其他人对于自己完成学位论文的时间长度是怎么说的？

环境支持

要完成如此巨大的工程需要什么样的工作环境呢？显然绝不会像找个地方吃早餐那么简单。因为此时你需要小心地选择工作空间。找一个安静而且不易被打

扰还能持续使用一段时间的地方吧,许多大学都为学生提供小的自习间。

我们的学生偶尔会为了完成博士学位论文而外租办公室。他们在办公室里只进行和学位论文有关的工作。这非常符合行为管理理论,在较少刺激的条件下控制行为,可以使人的精力更加集中。进行自我调控研究的研究者曾经有一段时间在谈论对于刺激的控制(Stuart,1977),有体重问题的人被告知只能在餐厅的桌子上放有白色餐桌垫的位置上进食;失眠症患者只能在床上入睡,而不能在其他地方睡觉,此时床作为一个控制刺激使睡觉变得更容易。斯金纳(B. F. Skinner)的房子里有一张专门用来写作的桌子。他每天都要在这张桌子上工作几个小时,而且不在上面做除写作以外其他事情(如付账单或者是写私人信件)。

假如把在一个在特定的地方进行写作当成一个控制刺激,那么与之相关的另外一个控制刺激就是固定的工作时间,如斯金纳每天早上在5—7点之间进行写作。坚持每天在同一时间工作,将增加对暂时性刺激的控制。使外界的光亮和声音(不容易分心)和自己的生理状态始终保持一致。尽管有时这样的"固执"是狭隘思想的魔鬼,但它却是你成功完成学位论文的必要因素。

获得家庭和爱人对你的支持就更容易获得这种"固执"性,不要忽视这些重要的支持。向他们坦白你需要持续一段时间像"鼹鼠"一样地生活,同时,安排出一些专门陪伴他们的时间。如果孩子和其他家庭成员知道你会保证安排专门的时间出来陪他们,那他们一般也不会去打扰你的工作。

计算机的使用和技术

你使用计算机和打字的技术如何(题目22和23)? 这些技能真的很重要吗? 在这个信息时代很难相信还有人会问到这样的问题,答案当然是:"非常重要。"虽然使用计算机的能力不是必需的,即使是在今天我们也遇到过不会打字的人成功完成了他们的学位论文的情况,但是假如你不会打字,那么你就需要准备出打字员帮你录入论文的时间。也许考虑到这一点,可能会迫使你投入精力成为一个会使用计算机的"现代人"——虽然能够使用计算机并不是强制性的要求,但却是现代社会必不可少的技能之一。

时间管理

第3章讲述了很多关于时间管理(题目25)技术的问题,这是我们人人都可以获得提高的一个方面。就在前面的部分,我们还讨论了一些通过控制刺激达到有效进行时间管理的方法。对于绝大多数喜欢拖延工作的人来说,如何有效地管理时间是他们完成学位论文时所要面对的主要挑战。现在是开始养成全新工作习惯的时候了。

社交技巧

题目26和27是关于在进行研究时如何与他人相处的问题。人际与社交技巧涉及能否和研究中的相关人员进行有效的互动。当然最重要的是和你的指导教师进行交流互动。你具有和导师在工作中进行有效互动的技巧吗? 你愿意为一些鸡毛蒜皮的零碎批评而把论文重写无数次吗? 通常学生们认为多次的审阅并不能使一篇科学文章更加清晰、连贯和翔实。但是据说艾萨克·牛顿爵士就很推崇多次复查的重要性。他非常注意复查自己的文章,经常是写作—修改—更正—重抄一遍,

然后再重新继续这个过程。“他的《数学原理》一书中的**一般评估原理**这部分内容曾经修改了8次之多”(Koyre,1965,p.262)。你也准备好这样做了吗?

你是否知道如何拒绝一些正面的建议?还是会对这些修改建议采取典型的方式进行辩护?你知晓答辩委员会成员之间的关系吗?你会选择那些不会干涉论文写作的人作为答辩委员会的成员吗?你能了解答辩会的风格以便在碰到一些问题时应对自如吗?如果你的研究需要校外人员的参与,你有这样的人际资源吗?你知道如何自由地支配和协调他们的方法吗?

如果你的人际技巧有一点粗糙,就把你有问题的方面一一列出来。你可以招募一个公关助手来帮助你和必要的人员进行协调,这个助手可以帮你完成大部分的工作。碰到必须亲自解决的问题,就多听取一些遇到过相同情况的同学和老师的建议。和你的朋友演练一下上述情境,然后让他们对你的表现做出反馈。

关于学位论文的正式/非正式规则

所有的学校都对学位论文的管理程序有一个正式的和非正式的规则。除去其他一些事项外,一致的地方有:需要填写一些表格、缴纳一定的费用、答辩的最后期限以及论文应该遵循的格式等。例如,许多学校会通知什么时候举行博士论文答辩会,答辩会都对哪些人开放。要想准确地通知一个答辩会的举行,一定要写明举行的时间和地点,你所在的学校也是这样吗?一些学校会给答辩委员会成员发一些时间表,他们要填写是否同意以及确实收到了学位论文的副本,你的学校是否也这样做?答辩委员会成员会在开题报告会举行的两个星期之前重新审查开题内容,你所在的学院也是这样安排的吗?早点弄清这些正式的规定,向你所在学院的教务秘书要一份关于这方面要求的书面材料做参考吧。

当然每个学院还有无数的非正式规则。时刻提醒自己遵守这些规则,可以让生活更加顺利。例如,你应该在什么时候邀请一名教师主持你的答辩委员会比较合适呢?在联系他/她之前,你需要向他/她提交一篇比较完善的研究思路和文献报告吗?如果你已经和几名教师讨论过一些研究想法,然后也很喜欢其中一个人的想法,但是却想和另外一个老师合作,该怎么办呢?采用一个教师的观点然后却选另一个人来做导师是否合适呢?提交导师审阅的文章完成度应该有多高?提交给答辩委员会的呢?不能安排开题报告的时间段是什么时候(例如每学期第一周和最后一周或是暑假里)?最好早点弄清楚这些非正式的规则,来个未雨绸缪。和快要完成学位论文的同学探讨一下这些问题,和导师还有潜在的答辩委员会成员交流一下这些规则。这些都是你不可能在教务秘书那里得到的信息。尽管有时非正式规则和正式规则一样重要,但是如果你不开口询问的话,很少有教师或者高年级学生会主动告诉你这些内容。

研究经费和资源

做研究不仅要花费时间,同样也需要花费资金。学校和教师为学生们的研究提供了大量的、从全额支援到完全自费不等的经费支持。如果教师有大量的研究经费,而学生的项目又正好在老师的研究课题之下,一般会得到全额的资助。

研究生支持协会(ASGS)通过电子邮件的方式收集了来自不同专业的85名学生的数据(“在失去资助的时候挽救你的论文”,1995)。大多数学生都能获得一定

的经费资助,因此可以选择进入那些经费资源丰富的学院。大部分学生的花费报告显示,一般的开销主要集中在影音材料、纸张、计算机和打印机的使用等方面。其他额外的开销要根据你的研究计划来确定,如支付给被试和研究助手的报酬,购买量表、软件的使用权,租用仪器,邮寄以及广告费等开支。也可能还包括外出的差旅费、调用档案的开销、申请学位的费用等。

那么应该如何支付这些花销呢?也许目前你还不能确定论文的题目,但现在也是做些计划的时候了。一些教师和院系有资助学生的项目,也有一些学校有通过竞争取得的学位论文资助款项。许多国家及组织,例如 APA 和国家健康研究院,也提供相应的资助和奖金,其中的一些直接发给特定专业的学生。尤其是在研究主题已经确定的时候,大学的资助和联系办公室可以给你提供不少有用的资源。网络则是另一种获取资助信息的来源(例如 APA 对心理学教育方面提供的资助(http://www.apa.org/students/funding.html)。最后,那些拿到过资助的学生和他们的导师也是有用的信息来源。现在,该对那些高年级的学生是如何控制他们研究的花费有一个总体的印象了。如果觉得获得资助的机会遥不可及,那就省下点儿喝咖啡的钱以备不时之需吧。

检查自己的认知状态

在进入论文写作的总体准备之前,弄清一些可能会阻碍你工作进程的情况将非常有用。马奥尼和马奥尼(Mahoney, & Mahoney, 1976)将其归为对个人的认知,即弄清自己的非理性想法。第一个非理性的想法是估计完成工作的时间。在读了这么多东西之后,你可能会对自己说:如果现在就开始工作的话,6 个月内就可以完成一篇学位论文;你也可能认为 5 天之内就能完成文字的编辑工作;或者认为自己可以从两所小学里得到所有的被试;或是和小学老师联系拿到研究许可只需要两周的时间等。牛顿的主要管理研究(Newtonian principal governing research)发现这些估计基本上都是错误的。实际上,大部分行为研究的结果也确实是这样的。无论我们什么时候进行估价,都可能经常出偏差,因此“比例运算”非常有帮助,即把你估计的时间乘以 3 倍就比较接近现实了。

我们容易产生的另外一个与上面类似的非理性想法是认为所有事情都会顺利进行,但事实并不会这样。我们应该为突发事件做好充分的准备,这样不管发生任何事,都是可以应对的。所以需要使用估计时间 3 倍的长度来完成所有事情,这就可以应付像被试没有按时到达、仪器的故障和数据的丢失、答辩委员会主席换了工作或是离开了学校等情况了。你准备好面对这些事了吗?注意这些不确定事件的突发性,并且准备一个后备计划吧。如果一个影响非常大的事件极有可能发生(如你的答辩委员会主席辞职了),以至于你不能够完成计划,那就必定需要同时准备一个可以替代的方案。

另外一个不理智想法是自己做出的研究必须具有权威性。我们没说有抱负有什么不对,问题是抱着一个要获“诺贝尔奖”情结是不对的,如果认为自己需要做出权威的研究,那么就永远都会觉得没足够的文献、研究问题的框架存在缺陷、不能做出合适的假设、没有选择最恰当的设计或是无法控制所有的重要变量等。时刻记住

知识的增加是一个递增的过程,利用一个研究就能做出科学发现的突破和研究范式革命(Kuhn,1970)的情况是很罕见的。

方法论的“完美神话”会阻碍研究想法的产生。每一个研究都有不成功的地方,如果不对相关的研究工作进行详细考察可能就不会发现这些错误,但是无论如何它们仍然存在。从最小的方面来说研究就存在权衡问题,如最常见的研究的内外部效度(一方面要科学地进行控制,另一方面还要在现实中推广适用)问题。想要让你的研究与众不同几乎是不现实的,我们不是说你不应该为变得杰出而努力,但是专注于使用完美的设计进行无懈可击的研究无异于天方夜谭。

如果你仍然觉得自己的研究应该具有完美的设计,是一个权威的研究,那就应该好好分析一下产生这些想法的原因。你能从中获得什么? 它们能使你在无法完成论文的时候免受自己和他人的批评吗? 因为想做一个完美的研究而拒绝一点点的缺陷到底是为什么呢? 是不是这样就可以打着保持研究严谨性的旗号而延期毕业了? 这样你也就可以留在学校里,不用到冷漠和残酷的现实世界中寻找工作和独立生存了?

本章的主要内容是关于进行学位论文写作的准备工作。涉及一些帮助大家顺利完成论文的相关技术和知识,其中的一些内容可以使你或多或少地得到一些改变。这里我们讨论了完成学位论文时产生的不理性的想法,剩下的一些问题则是关于学生需要进行很好准备的内容(测量和统计),这些知识可以通过进行额外的课程学习和训练来完成。最后是一些需要精通的基本技能(如语言的连贯性,语法的正确使用等)。如果你能够坦率、诚实地进行自我评价,那么本章的指导对你会有很大的帮助。坦率的回答研究清单上的问题,承认有些方面有待提高并不丢人,最好现在就发现这些不足并作出改进,而不是等以后通过“自卫”的方式来解决。

我们希望你能够认真地进行自我评估,学位论文是一个包括了许多技术的复杂研究项目。记住:你不需要完美地完成一篇学位论文,在缺乏经验或专长的工作领域这是很普遍的情况,我们都曾经以某种方式在这一过程中一些事情做得不够理想,但是这并不影响我们今后的成功。如果过去你在学术领域中有所建树,那么就在自己薄弱的方面应用过去的一些经验技巧,这样也能帮助你获得成功。

研究清单中的题目 25 是关于时间管理技能的,这些内容非常有用。但是即使是像作者这样经验丰富的人在完成复杂科研项目的时候,也会遇到这样那样的问题。因此,第 3 章的主题就是时间和问题的管理。下一章将为如何解决论文写作过程中出现的问题,使你能够继续顺利地进行工作提出许多有用的建议。

补充资源

关于思考和写作学位论文的软件的指导

APA-Style Helper 5.0:从 http://www.apastyle.org/stylehelper/ver5 获取

Error Detector:从 elc.polyu.edu.hk/CILL/errordetector.htm 获取

StyleEase for APA style:从 Stylewriter-USA.com 获取

StyleWriter™:从 Stywriter-USA. com 获取。更多关于 StyleWriter™ 的信息,参见 http://www. editorsoftware. com/。你可以从 http://www. editorsoftware. com/demonstrations/stylewriter-writing-software-demonstration. html 或者是 http://www. editorsoftware. com/reviews/RAA/download/stylewriter-trial-form. html 获得 30 天的试用版。

书籍和网站

American Psychological Association. (2001). *Publication manual of the American Psychological Association*(5th ed.). Washington, DC: Author.

Association for Support of Graduate Students. (2005). 网站 2005 年 10 月 31 日开放 http://www. asgs. org/

Darley, J. M., Roediger, H. L., & Zanna, M. P. (Eds.). (2003). *The complete academic: A practical guide for the beginning social scientist*. Washington, DC: American Psychological Association.

Krathwohl, D. R. (1988). *How to prepare a research proposal: Suggestions for those seeking funds for behavioral science research*(3rd ed.). Syracuse, NY: School of Education, Syracuse University.

Meltzoff, J. (1998). *Critical thinking about research: Psychology and related fields*. Washington, DC: American Psychlgical Association.

思考与练习

评估你的准备情况

- ☐ 完成研究清单
- ☐ 对测验结果的解释
  - —写作技能
  - —方法学技能
  - —环境支持
  - —计算机技能
  - —时间管理技能
  - —社交技能
- ☐ 检查认知生态性

3 时间与问题管理

在第 2 章中,我们已经提醒你要预想到在完成学位论文的过程中可能会遇到的突发事件。本章的主要内容包括在完成学位论文的过程中最常遇到的问题,以及如何避免和克服它们的建议。本书的第 10 章会帮助你规划实施研究的细节。这两个章节的目标是帮助你对研究进行提前的预期和计划,这样可以避免意外的发生。

本章要讨论的第一个问题是如何有效地管理时间。爱因斯坦(1974)把时间定义为事件发展序列的逻辑延伸。因此有效地使用时间就是有效地管理事情的发展。但是如何才能做到"有效"呢?"有效"意味着"获得理想的结果"。有效地管理事情也就是合理地安排事件以获得期望的结果的过程。因为生活中的大多数活动(尤其是研究项目)会涉及一系列的不同事件,所以对你来说有效地管理时间就意味着有效地操纵这些事件并且顺利完成学位论文。

开始的目标

想成功地完成一个研究计划,需要以一个目标作为开端。你需要仔细地制订一个包含确切内容的目标。最重要的是,这个目标要包含“你要完成**什么**”并且“在**什么时候**能够完成”这样的内容。为了得到最佳效果,前一句中的两个“什么”应该指的是可预测的行为。例如“在6月30日完成我的计划,____”,或者是“在5月1日之前获得小学的许可,____”。

设定目标时最困难的部分是确定“什么时间”这个元素。有时这个“时间”由你自己决定。不管你是否能完成学位论文,对学生获得资金资助的期限和要求是固定的。例如,对外国留学生发放的政府奖学金规定:毕业之后学生必须回国服务一定的年限。显然,这种时限要求比仅仅失去经济资助这样的条件要明确得多,也比开放性的论文完成日期要精确得多。如果你要在一个确定的时间段内完成学位论文,那么能够仔细地计划和有效地分配你的时间是非常重要的。

评估你需要的时间

把你的学位论文当成一个工作项目来完成。不管多大的项目,都包含着数量不等的子项目或者是小步骤,逐个完成它们才能够最终完结整个项目。其实你可以通过仔细地鉴别和规划这些小步骤来从容地控制自己的研究。那么到底应该如何来明确这些小步骤呢?

首先,坐下来思考一下你所知道的关于论文的全部要求。记住,正式的和非正式的规则都要考虑到。如果你不知道这些,那么现在就赶快去弄清楚。把你需要做的事情列成一张表。为了最后能比较轻松地完成,最好写在3英寸×5英寸的卡片上面。一件事写在一张卡片上,以动词形式来开头,例如,“招募被试”“寻找工作场所”等。学位论文最终版的复印件应该保管在学校的图书馆里,把这个事件标成“把____上交到图书馆”。你也许会问为什么每个项目都要用动词开始?这是因为动词涉及的是动作,而且对于谁来执行这个动作也非常清楚。比较“寻找被试”和“被试需要被寻找到”这两个句子。第一句是一个非常清楚的指示,而第二个仅仅是一种宣告。所以第二句根本没有什么作用。人会按照说话的方式来行动,果断的语言也伴随着果断的行动。比较“试着招募被试”和“招募被试”这两句的有力程度。你在邀请别人参加派对时,愿意听到他们说:“听起来很有趣!我尽量去”吗?你觉得他们真的会出现吗?所以一定要明确你的目的,然后果断地完成。

我们通常要求学生们独立认真地制作一份目标清单。这份清单应该包括从什么时候开始工作到什么时候能完成一篇学位论文要做的、所有必要事件的步骤。很多学生发现清单可以把需要完成的任务具体化,并且对他们非常有帮助。尽管我们提倡学生们按照“小步子”原则去制订些小的和容易完成的步骤,但是他们通常会自由地制订清单中的各个步骤。我们不会对事件序列的数量做出规定。一些人喜欢将工作进行大块的划分,例如,“写作论文”。其他人喜欢小步地进行工作,例如

“列出主题提纲”,“写出前5页内容”和“把第一稿交给导师看”等。

排列清单中步骤的顺序有两种方法。从后向前的逆序式(如“上交到图书馆”)和从前往后的顺序式。如果你有一个明确的完成论文的期限,那么最好的规划方法是采用逆序式。如果你没有一个明确的完成日期,则可以比较随意地使用顺序式方法。列出从开始写作论文到结束整个过程中所有需要完成的步骤,然后给每一个步骤加上一个完成限期。如果你觉得为整个学位论文设定完成目标的任务过于巨大,可以列出一些中间目标(如,“获得对开题报告的批准”),然后再设定其他的完成步骤。

有些人的清单目标只有6步,也有人设定的目标多达46步。一些人创造性地把他们的清单做成了包含各种替换事件和基于结果进行多重选择的流程图。例如,一个学生清单中的可替换事件是“成功的答辩”。“如果到时候得到的结果是‘是’,那么就和妻子、家人去欧洲旅游。如果得到的是‘否’,则再从头开始一切”。在这一过程中有一点点的放松可能会给你带来很大的帮助。

在按照清单工作时,要把每一个小项都当成一个目标来完成。接下来把完成目标的具体步骤写清楚,并且标明需要的行动和完成的时间。如果你从完成时间的期限开始设定项目,那么使用逆序的方法非常有效。

编排工作

一旦你已经确定和分配完论文写作过程中每一个步骤所需的时间,就会发现把这些内容进行可视化是非常有用的。这里有一些你需要的东西,包括简单的Gantt图到更加复杂的PERT(Program Evaluation & Review Technique)分析。这些工具可以给你提供一个计划框架。例如,Gantt图可以列出一个可视的时间进程。很多的时间规划的软件(例如MS Project™)也会对你起到帮助作用,这些软件能够进行从简单的时间图制作一直到成本评估、条件可能性问题分析等操作。除非你是个项目时间安排天才,或者是非常有兴趣深入研究时间管理技能,不然这些软件已经足够你用来制作出合适的Gantt图或是时间图了。(空白的Gantt图可以从专门的商店买到。)

专栏3.1是一个研究项目的Gantt图,左边是所有的行动步骤。那么就按计划开始执行吧。专栏上面的部分是完成这些步骤所需要的时间。箭头代表一个行动或者是步骤开始和结束的时间。发生一次的行为用×表示;有时也用△来表示。从行动1中看出,儿童评估从第一周的第一天开始,一直持续整周。行动2中,确定目标儿童和伙伴,也从第一周开始持续整周。行动3,随机分组,只发生一次,在第一周的末尾。行动9,进行直接观察,在第4周和第8周开始实施。很显然一些事件需要在另一些之前完成,一些是同时完成的,还有一些可以重叠着进行。

专栏 3.1

儿童社会技能研究进程图

行动 \ 周	1	2	3	4	5	6	7	8	9
1. 评估儿童	→								
2. 确定目标儿童和玩伴	→								
3. 随机分组	×								
4. 访谈匹配组的儿童和玩伴		→							
5. 发展模板		×							
6. 和目标儿童所选择的对象谈话			→						
7. 为匹配组设定目标行为		×							
8. 设计处理计划		×							
9. 进行直接观察				→				→	
10. 实施处理				—	—	—	—	→	
11. 重复课堂评估									→

对这些步骤和它们之间这种暂时关系的视觉呈现可以使你更加明确计划的完成情况。它也显示了一些互相依赖的事件间的关系，使得整个工作变得更加清晰。而且这份表也可以作为向别人描述的工具，在你与希望招募的研究对象的代理人交流时会更加有用。这个表不仅解释了你需要做什么，也向答辩委员会表明你已经全面地考虑了所有问题，或者说是基本考虑了所有问题。

简单的计算机时间表安排软件也可以即时地修改你的研究计划。如果你熟悉计算机中的时间表软件，就会知道当一个单独的数据改变后，其他的相关值都会被自动重新计算。例如你因为一些意外而准备在研究过程中间休假一周。举个例子来说，也许是你需要跟随自己的阿姨去夏威夷度假，因为你是唯一可以陪同她的人（当然，她是这么认为的），这会对你的整个研究产生什么影响呢？只是需要简单地把所有日期都延后一周吗？还是因为计划中存在一些与其他事物相互牵制的同时性的行动或者条件而变得更加复杂？时间表的安排并不是一成不变的。在你开始项目的时候就得改变一下自己已有的逻辑习惯。你也可能在掌握编排工作这个技术之前不知道如何估计时间，进而导致你论文的最后完成期限的改变。而时间表软件可以向你直观地呈现出这些变化。

用计划时间表来解放自己

此刻,你一定会问,为什么对日程安排、日期和表格等细节如此小题大做。这些良好的计划能避免出现问题以使你继续自己的研究吗?答案是肯定的。一丁点儿的预先计划就可以避免后面出现大量的问题。在军队中,一个中士曾经用5个P告诫新兵,那就是:**通过预先计划来预防糟糕的后果**(prior planning prevents poor performance)。花一点时间去确定研究项目、排序和时间安排可以为以后节约大量的时间,并避免遇到麻烦。

另外,按照时间表来完成你的学位论文会帮助你把它变成日常生活中的一个基本组成部分。它可以使你明白自己是否有时间按部就班地进行学位论文写作,或者需要停止其他的活动来专心致志地完成它。以我们的经验来说,那些每天严格按照时间表执行的学生们可以较早地完成学位论文,而且也可以采取较少的规划步骤。还记得龟兔赛跑的故事吗?知道最后是谁赢得了比赛吗?在做学位论文的时候也是如此。

制作一份时间表和按照此表严格地执行同样重要。如果你勤勤恳恳地确定了时间表的所有步骤,就已经完成了计划中最难的部分,剩下要做的就是照计划完成时间表上的每一步。这也是我们为什么鼓励学生把计划设成许多小步骤的原因。最好是让它们小到可以用一天或者是一周的时间来完成。如果你这样做了,就可以把这些记到日程本或是掌上电脑上,每一天写上需要完成的步骤。

明确你在行动中或是相关的条件上希望获得的结果。对于大多数人来说,"读5篇文章"比"用2个小时来阅读"更好,当然"阅读并总结5篇文章"就更好了。一些学生告诉我们,他们觉得时序目标比确定的产品目标更有用,可能是因为他们已经在使用时序计划了。关键是要保证日常工作按照时间表进行,并且一步一步地完成。

一旦所有的步骤都记录到你的日程上,就不用再焦虑了。你所要做的工作就是按照日期的安排完成每一小步。就像一个老笑话一样:如何吃掉一头大象?一口一口地吃。你就要这样完成所有的研究任务。实际上,在我们撰写本书第1版的时候,另外一个作者就制订了每周写5页书的计划并且严格执行,她在1个月内按时完成了自己需要写作的章节。

我们曾经有一个学生在每学期的开始就安排好所有的课程、研究、作业以及课堂项目的计划,甚至每一天该干什么都安排好了。当其他学生问他如何能够忍受执行这样一个详细的时间表时,他回答说事实正相反,他觉得整个程序非常自由,这样就不必花费时间去思考如何完成所有的工作或是担心自己是否有足够的时间来完成工作了。他可以有更多的时间用来娱乐和消遣(当然是有计划的)。通过制订计划来完成所有工作的方法不是对所有人都适用,但要记住:当你知道要做什么并且按计划执行的时候,你就可以没有负罪感地去享受其他的时间了。

最后要记住的一点就是要为你的亲人好友在时间表上保留足够的时间。如果在一个特定的期限内你和那些对你很重要的人在一起共处的时间会被缩短,一定要事先和他们谈一谈。他们需要预先得到通知并且需要在这段时间内安排各自的计

划。如果可以的话，把他们也加入到你的时间表中来。这样他们就不会因为被排除在外而感到气愤了。平衡工作和个人生活从来都不是一件容易的事，确保在考虑到健康、朋友和家人的同时还要按时完成研究是一个巨大的挑战。

我们建议早点使用时间表并制订工作计划以避免遇到意外的问题。因为这个时候你更加容易知晓在完成论文过程中存在的一些其他需求。例如，没有专门给学位论文申请人准备的整洁的房间。在做学位论文时延误(P)和逃避(A)也很常见，具体内容可以在《精神障碍诊断和统计手册》①中找到。

我们都会逃避一些事情。然而，在极端情况下，P和A可能导致你不能完成学位论文。因此，如果你被P和A控制了，你就成了ABD(All-But-Dissertation，什么都做了，就是没有完成论文)。为了避免这种结果，把下面的内容用大字体打印出来挂在墙上。把它作为与研究同步的日常提醒。我们的一些学生会把这些记在心里并且把他们的目标贴在冰箱或是电脑上来提醒自己注意。

P和A导致ABD

如果按照本章前面给出的建议去做，你就已经掌控了P和A。如果你确定了足够细致的步骤、适当的安排顺序，并且把它们记在你的日程本或是时间表上，就已经准备好开始完成自己的学位论文了。你有没有过这样的经验：轻易就能完成的事情会促使你有动力去做其他事情，学位论文就是这样完成的。每天完成一小步，释放出来的激励能量让你继续完成更多的步骤；持续进行以产生更多的激励能量等。一个循环就这样建立起来，并且可以促进论文的完成。

对于主要的研究项目而言，时间主要耽搁在论文写作、确定题目、开题以及撰写学位论文的最后一版上。如果到目前为止你一直按照我们的建议执行，但是还是存在延误的话，也许你用了以下的一些行为动词欺骗了自己。

1. “我不能开始。这个项目实在是太巨大了。”感到项目巨大是因为你没有把任务分成小的步骤。想想你今天能够完成的目标：读5篇文章？打出论文的目录？还是列出论文主要部分的提纲？巨大的任务可以通过将其分成一些小步骤来完成。如果你不能面对自己制订的计划，那就做点看起来更可行的事情。完成任务可以带给你动力去处理其他的事情。一定要做一些事情。

2. “如果没有足够的时间，我就没办法进行研究，而且到下一个暑假之前我都没有时间。”这是一个在下一个暑期之前什么都不用做的、非常有创意的借口。事实上，即使你在有大块时间可供支配时工作会更有效率，但是在缺少大块时间的时候你也可以完成学位论文。另外，难道你就不能每周抽出至少3个时间段来做你的研究吗？例如本书的一个作者就在过去的10年中每天早上都用来进行学术写作。这个时间就和上课的时间(你不会因为要做作业而不去上课吧？也不会为了打扫公寓、购买食物而逃课对吗？)一样不可侵犯。过去的几年中，她的很多文章和书就是在这个时间段里写成的。就像我们反复强调的那样，人们通过持续稳定的工作来完成学位论文。

如何更好地利用时间也很重要(见Yates，1982)。我们的学生在如何更好地工作、什么时间、哪个时间范围、单一还是同时地进行多个工作等方面差别很大。最好

---

① *Diagnostic and Statistcal Manual of Mental Disorders*(DSM, of the American Psychiatric Association, 1994)

在头脑最清醒的时候工作,如果你的最佳工作时间是早上,就不要在9—11点之间做你的学位论文!在你累了或者并非最佳认知状态的时候进行些难度较低的任务吧。

3."我知道得太少(读的文献太少、研究解决的问题太少或者是没有足够的假设)以至于写不出什么来。"这是完美主义者的思维变式:"我没办法完成不完美的作品。"如果你执迷于此,那将什么也写不出来。论文的草稿是不可能完美的。只要写就行了,要知道它还没有完成呢。然后,再进行修改,完善第一稿。别忘了回忆一下第2章中的"语法检查/编辑软件"的有关内容。

更成问题的是自我质疑:"我不够聪明,永远不能让我的导师满意,我是个'骗子'(Rudestam & Newton,2001)。"这种自言自语会动摇你的自信然后使你产生不必要的情绪行为和自我指向的负面想法。有必要区分清楚学位论文(你要面对的现实)和对自己的评价(你自己的假设和哲学观)是两回事。是的,学位论文是个挑战,它很可怕,它是通向未来的一扇门,甚至更可怕!是的,它是一个漫长和艰苦的任务,要求你不断鞭策自己努力工作。但是,你辛苦完成的草稿可能会给自己带来"建设性"的建议,这样你可以选择战胜自己。是的,有时你会迷失,但这只是暂时性的迷失。这些并不意味着你笨,是一个骗子,或是一个上不了台面的人。绝大多数做过学位论文的人都要经历这些事情。在完成一个具有挑战性的任务过程中这些感受是很常见的。如果你觉得这个过程很难,千万不要让自己被打倒。相反,拍拍自己的后背告诉自己:至少这些经历说明自己是一个正常人!

4."我只有在最后期限临近时才能工作。"一些人直到最后期限接近了才停止闲混,然后通宵达旦地赶出一篇文章、复习考试、完成开题报告,等等。当开始着手完成学位论文的时候,这些人发现他们处于麻烦之中,因为完成的过程包括许多小步骤和缺少准确截止时间的长期的工作目标。可能你已经意识到了,一个开题报告或是论文的最终稿不可能靠咖啡因或是安非他明维持一个周末的清醒就能完成。我们给截止期限工作者两个解决办法。首先,要给学位论文的完成时间设定一个真实有效的截止日期。也许你的研究生教学计划里就有这个期限。例如,一些教学计划规定,临床心理学学生在没有提交学位论文开题报告之前不可以申请实习。如果你还没有确定一个提交论文的准确日期,也许你的导师可以帮你创造条件。例如,如果你要申请一份学术工作,也许你和导师可以达成这样一个共识:当论文研究的一定步骤(收集数据或是数据分析)完成时,他可以帮你寄出推荐信等。确定其他的日期,如提交给大会的摘要数据或是在答辩会上报告你的结果的日期,也可以作为你工作的最后期限。

第二个建议是尽量打破到最后期限才工作的习惯。如果你想从事学术性工作,就需要减少拖延的毛病。所以为什么不趁现在开始改变呢?如果你已经像本章说明的那样设立并且按照日程表开始工作了,就可以改掉这个坏习惯。

通过普雷马克原理来安排你的时间表是一个有用的方法,普雷马克原理以发现者的名字来命名。一个简单的事实就是我们可以通过把比较容易的、喜欢的事情放在后面做来促进完成比较难的,不想去完成的事情。这是古老格言背后的行为理论"先工作,后享乐",先做一些完成学位论文的工作,然后再做那些你想做或者需要去做的事情(例如,购物或是睡觉)。在生活中建立"工作第一,玩乐第二"的规则,

这样你能够更轻松地完成论文。

在制作计划时间表时加入一些奖励是个很好的想法。在把开题报告最后一版上交给答辩委员会之后,计划一次和爱人一起的旅行吧。但是不要在草稿还没有完成之前就离开。如果你的奖励包含了和其他重要的人一起行动,那么他们也会一直支持你,你要做的就是让他们知道这一点。通过这种方法,你可以更好地平衡工作和生活,而不是等到论文全部完成之后再决定安排和他人一起做些有趣的事情。

最后,请你的家人、朋友一起来鼓励你。告诉他们你取得的成就,得到他们的祝贺。一些学生甚至组成了学位论文支持小组来使工作走上正轨。例如我最近的两名学生,每周见面并一起进行学位论文写作方面的工作。他们到图书馆,或者在学生公寓一起各自忙活各自的论文。另外一名学生加入了一个在线的学位论文支持小组,这个小组为她克服拖延习惯提供了许多很好的建议。尽管如此,也需要注意:互助小组是为督促你完成学位论文,而不是为了在不能完成论文时给你提供一个"好的借口"或博得同情! 我们还有一个建议就是在你的冰箱上贴一张每日(每周)工作进度表。大家可以根据上面的数据恰当地评估自己的工作情况。

在学位论文写作过程中产生的另一个问题就是如何排除一些对工作的干扰。许多人通过朋友、导师、同龄人的帮助或者使用本书的同类书籍(见本章后面的清单中的附加目录)来克服这些问题。"学位论文教练"和专门从事学生工作的心理医生也可以给出专业的指导。如果管理或者情绪问题严重地影响了你的学位论文进程,是否进行专业求助来协助你继续完成学位论文就是一个值得考虑的问题了。

就像我们一贯重复的,完成学位论文的关键是坚持不懈的工作。尽管我们提供了不少关于如何安排时间表和计划的建议,但是完成一篇学位论文仍然有很多不同的方式。最重要的是找到一个适合你的研究项目,利用你的知识、设备、前瞻性和计划性来完成它,这样你的工作就会步入正轨。

补充资源

Bolker, J. (1998). *Writing your dissertation in fifteen minutes a day*. New York: Henry Holt. [这本俏皮的书籍主要关注文章写作过程——考虑如何写的同时管理写作进程。]

Burka, J., & Yuen, L. (1983). Procrastination: *Why you do it, what to do about it*. Cambridge, MA: Da Capo Press. [受到学位论文教练麦克金尼的高度推荐,他是一名专门为毕业生和准备入学的学生提供帮助的临床心理学家。]

Gantt charts. (2005). Retrieved November 17, 2005, from http://www.ganttchart.com/index.html. [一个提供甘特图的总体描述和背景历史的网站。]

Lakin, A. (1974). *How to get control of your time and your life*. New York: Signet Books. [也是得到麦克金尼推荐的一本经典书籍。]

Tracy, B. (2004). *Time power: A proven system for getting more done in less time than you ever thought possible*. New York: AMACOM.

Yates, B. T. (1982). *Doing the dissertation: The nuts and bolts of psychological research*. Springfield, IL: Charles C. Thomas. [第2章的条目中提供了关于个人管理和时间管理以及对个人评估的很好的建议。]

思考与练习

管理时间和困难

- ☐ 设定目标
- ☐ 评估你需要的时间
  - —确定子目标和行动
  - —把大的任务分割成小步骤
  - —将子目标/行动/步骤排序
- ☐ 编时间表
- ☐ 把延迟和逃避降至最小化
  - —避免普遍的消极思维
  - —使用普雷马克原理
- ☐ 编排学位论文的时间,使其转化成你的工作动力

4

确定主题及导师

本章开始讨论学位论文中如何确定研究主题和怎样选择指导教师的问题。本书的章节顺序是按照绝大多数学生完成学位论文的程序来编排的。例如，我们在本章中讨论选题和挑选合作教师①的内容，在第6章中讨论如何写作学位论文中的文献回顾部分。假定你要完成一份综述型的论文，那么确定论文的选题就变得尤为重要了。这时可以在阅读完本章之后直接转入第6章，而不必一定按本书编排的顺序进行学习。本章我们重点讲解如何用语言精确地描述研究假设，学生们一般在仔细学习了文献回顾（第6章）和研究方法（第7章）的内容之后会更容易完善他们的研究假设。就像我们多次强调的那样，你需要寻找适合自己的研究顺序和方法。

现在，让我们从规划学位论文时要做出的两个重要决定入手：首先要确定学位论文研究的内容，其次要选择一位指导教师。因为这两者关系密切，所以放在这里一并进行考虑。

---

① 即挑选导师和答辩委员会主席，美国的答辩委员会主席一般由学生的导师担任。——译者注

选择一个研究领域

对很多研究生来讲,最困难的就是提出一个有研究价值的想法。这期间大家容易产生两个错误的观念:一是在第2章中提到的"诺贝尔奖效应";二是"本科研究论文"错误,即学生们认为他们产生的每个想法都是一个好的研究题目。尽管有些教授可以容忍这些论文的荒谬性,但是你必须对选择此类菜鸟题目持谨慎态度。

那为什么这些观念是错误的呢?首先,因为"诺贝尔奖效应"反映的是不切实际的宏伟构想;其次,"本科论文错误"反映的是不符合现实情况的渺小想法。这两种错误都反映了缺乏正确对待现实问题的情况。作为指导教师,通常我们更偏向于原谅那些受"诺贝尔奖效应"影响的学生。因为,去掉一些冗繁的研究设计要比加强一个羸弱的研究构想简单一些。而且,"诺贝尔奖效应"至少表明申请者愿意付出努力去工作!

没有太多实际研究经验的人经常出现上述两种错误。因此,除非博士生独自进行过一些研究工作,不然在完成学位论文时会比硕士生遇到的麻烦更多。如果你在过去的研究中得到的研究推论是可靠的,提出一个以此为基础的研究设想可能会收到不错的效果。如果你是硕士研究生,导师可能会给你指定一个研究领域。你可以将自己的研究建立在扩展那些已经发表的研究结果的基础上。如果你是一名博士研究生,那么研究想法就应该具有原创性,进行的研究也必须对你的专业领域做出相应的贡献。

一般来讲,要确定可行的研究项目需要完成两个步骤。首先,确定你的研究领域。例如,短时记忆、童年期侵犯、社会支持或是信息加工过程这样的大体方向。确定研究方向这个问题比较简单,而且在有人问你从事哪些研究的时候也有话可说。一旦你选定了一个领域,就进入提出研究问题和假设的阶段。这时你的研究问题和研究假设需要经过深思熟虑的锤炼而变得更加具体才行。大多数学生发现完成这一过程比选择一个研究领域要困难得多。

确定研究范围

让我们先谈一下策略问题。学生们经常问应该在什么时候开始选择论文题目。我们的答案是"越早越好"。少数幸运的人在本科的时候就已经确定了自己的研究方向。他们可能选修了自己感兴趣的研究领域的相关课程,可能做过老师的研究助手或是已经参与过其他研究生的部分研究工作,或者也可能因为喜欢的某个老师在从事某领域的研究工作而选择了该领域。对这些人来说,博士学位论文是他们过去做过的研究项目的延伸。然而,大多数人并不具备这样的研究经验。即使是有,也可能与现任导师的研究领域差别巨大。

如果你还没有开始选题的话,我建议你了解一下导师的研究兴趣。选出那些你感兴趣的问题并找出你可以从事的研究项目。当然也要知道老师们的兴趣所在并向他们多加请教。在你的研究开始之前,询问老师的研究方向以及他们是否可以给你提供一些参考资料,看看这些材料中是否有你想从事的研究。如果有,就直接开

始你的研究工作。从事科学研究工作需要很多复杂的技能，而直接获得的经验比单纯从书本上学习到的要重要得多。

最好的情况是让学生们在研究生开始的时候就开始准备，这样有助于进行后续的研究工作，使他们学到的每一门课都与其有关联。开始得越早，就越能更早地发现其间的联系。

在学习过程中，快速地记下那些灵光突现的想法是非常有用的。有些人每天还在电脑中记下自己对论文、书籍、研究项目的想法。

如果你开始准备论文的时间很早并且不喜欢现任导师或研究方向，那就换一个。在学业早期尽量和多个教师进行合作，熟悉他们各自的风格、工作方式、特点和其对学生的资助奖励并没有什么错。一些“高产”的学生在研究生生涯中和许多教师合作过。这样有助于他们从不同方面获得大量的研究经验，并且更容易产生很多可行的研究设想。如果你选择了这条道路，事先告诉你的导师，这样他们到时就不会对你的行动方式太过惊讶或是给你指定研究方向了。

如果你已经做了一段时间的研究生但是却没有什么研究经验的话，也不要感到气馁。记住，你首先要做的只是确定一个研究领域，而不是计划着要怎么完成整个研究计划。你选择的研究领域必须符合两个关键的标准：第一，你对它有足够的兴趣，愿意花费时间来阅读文献、撰写文章和分析相关数据；第二，找到一个对此也有兴趣的导师。

要确定一个研究领域，先想想你的课堂经验和专业知识。你对哪些研究主题感兴趣？在完成课堂研究作业的时候你一般会选择哪些资料来阅读？当你和其他老师、同学讨论问题的时候你最关注的是哪些学术主题？你最喜欢写哪类学术论文？你最近经常在作业中写到哪些问题？回答完这些问题你就可能已经开始展开学位论文的文献回顾工作了。在第 6 章中我们会介绍关于文献回顾的问题。有一些学生在选题的时候非常具有策略性，他们把自己当成一名心理学家，然后再来计划将要从事的工作，并且以此指导自己选题。

确定自己思考过、读到过、谈起过或计划要做的事是发现研究兴趣所在的好办法，你会发现不少感兴趣的领域。然后你要找出对这些领域也感兴趣的教师并和他们交流一下。尽管我们建议你在寻找导师之前最好先找到一个感兴趣的研究领域，但这不是绝对的。不少学生会逆向地进行这一过程，他们先寻找到喜欢的教师，然后再共同确定一个研究主题。如果你正在筹划一篇硕士论文，可以问问老师们是否有让学生参与的科研项目，这也是一个发现教师们各自特点的好机会。找出他们可能感兴趣的项目，然后给彼此留下一个好印象。总而言之，最好选择一个能获得老师的帮助却又不是特别热门的领域，这绝对好过只挑选一个自己觉得有趣的奇怪领域。

避免独自进行

一些学生，特别是有些博士生，总是认为他们应该独自展开研究。所以会花费数月的时间完成一个研究设想和开题报告，然后寻找一位老师来兜售他们的“产品”。耶茨（Yates，1982）说他自己就有过相同的经历，并且当导师因为研究内容的

不可行性而拒绝他的时候感到强烈的崩溃和绝望。碰到这种情况时可以换一个能够赞同你想法的老师，或者重新开始另外一个研究构想。后者在研究生院里可不是什么时髦的事情，所以应该谨慎对待这一问题。

确实，博士学位论文是对你是否具有进行独立研究工作能力的一项检验。不管是在产生还是在完善研究想法的层面上，硕士学位论文的作用与此是不同的。事实上，独自进行工作很容易产生出不可行的设想。另一个频繁出现的问题是，不能找到一个在这方面具有兴趣和专长的老师在整个研究过程中引导你。我们知道曾经就有学生带着自己的想法拜访了院里所有的老师却发现找不到一个能够指导他的人，缺乏良好的推销技能也使完成学位论文的工作变成了一项严峻挑战。鉴于上述两点原因，我们强烈建议你不要独自进行自己的论文研究。通过和教师合作的方式完善你的想法，并经常和老师保持联系。如果老师认同你的观点，你们会保持非常良好的合作关系。

选择导师

为你的学位论文挑选一位导师与论文选题一样至关重要。一个好的导师需要了解你的研究领域，不但为你的工作提供特定的评论，而且能够指导你并在需要的时候提醒你应该怎样做。相反，一个不合适的导师不仅做不到这些，还可能在你精心准备学位论文的过程中使你感到困惑不解。

在把某个教师变成潜在的导师之前，我们需要考虑一下主持答辩会的规则都有哪些。尽管每个学校和导师的风格都不尽相同，但是其中的一些共同点还是存在的。首先，导师要帮助学生们完善研究想法和实验方法。第二，他们要对研究项目进行第一轮的评估工作，因此他们要对学位论文的每个部分进行多次审阅。第三，要获得他们许可，学生才能将文章提交给答辩委员会的其他成员。

就硕士学位论文来说，大多数的导师和学生的联系都很紧密。而在完成博士论文的阶段，导师们的表现就不尽相同了。有些人希望完全参与到学位论文工作的整个过程中去，而另外一些导师则希望只有在你遇到问题和快要完成论文时才向他们汇报。如果你喜欢独立工作，那么就选择一个能"放手"的导师，不然就选一位更主动的导师。在选择导师的时候你要考虑三个问题：①如何与导师合作。②他们对你的研究领域具有丰富的经验吗？③能否指导你顺利完成学位论文？如果你选得好，你会和导师合作愉快，并且得到他们具体而又省时的专门指导。与一个好的导师合作，就像看一个好牙医：虽然整个过程可能并不是十分愉快或是没有丝毫痛苦的，但也绝不像你想象的那么糟糕，最后的结果就是物超所值。

学院里一般对导师和答辩委员会成员资格有专门的规定：除非在特殊情况下，不然导师必须具有博士学位。一些学院只允许那些近年来发表文章数目足够多的或是从事一定年限教学工作的人做导师。制定这些规则是因为没有经验的教师或是那些没有获得相应学位的人，不能保证让一篇博士论文在学术上拥有足够的广度、深度以及使用了恰当的研究技术。符合担任硕士生导师资格的规定不一定适用于博士生导师。进一步来说，一些院校的学生要想和本院以外的教师或者其他领域的老师合作必须要拿到特别的许可。只有在特殊情况下（例如，当有一名教师同意

成为名义上的导师或者是合作性导师时）才允许非教职人员担任导师。遵循这些规则可以帮助你避免与不适合成为导师的人合作。另外，在找到符合这些规则并答应成为你的导师的人选之后，别忘了看看还需要和导师、答辩委员会成员签署哪些表格，这些正式文件会让他们在将来遵守和你的约定。

考虑一下导师—学生的关系

一篇学位论文是导师和学生共同合作的产物。当把一篇学位论文提交给《国际学位文摘》的时候，导师和学生的名字都会被列出。因此那些在国际上和国内知名的教师会格外谨慎地对待论文的指导工作。担任一个学位论文答辩委员会的主席可不像大多数教师想象的那么轻松。你和导师需要建立良好的合作关系以确保能够成功完成论文。特别是在你只有很少的研究经验并且需要很多指导（经常出现在硕士论文阶段）的情况下更是如此。评估自己如何同导师合作的第一步是想想自己希望从导师那里获得哪些帮助，这是使导师的指导或工作方式能够行之有效的前提。所以，你想从导师那里得到什么呢？你需要有人从研究领域到具体论题（如果导师也对这个问题感兴趣的话）进行全程指导吗？还是喜欢找一个可以让你自由发挥的导师？你是与具有高度组织化（例如设定了截止期限和与学生定期会面）工作风格的导师合作还是同"放任型"的导师进行合作效果更好呢？你想与一个温和的指导型的导师合作还是与一个具有其他特质（例如经验和专业性）的导师合作呢？简而言之，你会发现老师的个人风格和你的工作风格相契合的程度是个重要问题。为此，你需要认真考虑和对下列事项作出决定：打算采用哪种文章结构？为研究主题的完善进行多少工作投入？以及你喜欢哪种风格的导师？

考虑潜在导师的经验

一些学生和教师认为"在研究主题上，导师必须比学生知道得多"。也有些人不同意这种观点，他们认为导师的责任是对论文中的研究方法和有争论的逻辑问题做出批判，不一定是你那个主题领域的专家，因为好的方法学理论素养和分析技术在哪个领域都适用。我们同意后者的观点，我们甚至认为在导师对某一领域了解甚微的情况下，有时也可以指导学生写出好的学位论文来。除此之外，我们相信对学生和老师都有利的情况就是学生能在自己要从事的研究领域找到一个有经验的导师。即，我们建议学生能够和一位在这方面也有兴趣的老师一起完善具体的研究问题。这对那些没有什么研究经验并且需要更多指导的硕士生来说非常重要。

我们强烈建议与一位熟悉你研究领域的导师合作，在我们看来，这样能花费更少的时间更轻松地产出更好的作品。那些在某一领域中工作了许多年的教师已经很好地掌握了相关领域的各种文献。他们知道哪些研究已经做过了、哪些还没有做以及哪些是现在刚刚完成的。这些知识可以防止你陷入窘境，例如，他知道你花费数月时间设计出来的实验于不久之前已经在《行为和社会心理学》上发表了。进一步说，这样的导师也可以指导你做出更好的研究。

同样重要的是，博学的教师了解研究领域中的许多细节。他们知道如何有效地招募被试、设计实验以及实验程序的一般缺陷在哪里、实验范式都有哪些等。这样就不会使你因为没有足够的时间而只完成了一篇硕士水平的论文，也不可能因为

(你或是导师)对某个领域的无知而设计出一个有致命缺陷的研究。一个合格的导师可以直接指导你阅读相关的论文和书籍。例如我们的一个学生说要去图书馆查资料。后来她发现导师有满满两个抽屉的相关文献,而且只用一天的时间就复印到了所有需要的文章,她为此节省了许多搜索文章和下载文献的时间。

另一个在导师从事的领域选题的原因是:这样他会在你的研究项目上投入更多的精力。那些关注你研究的导师会花更多的时间并更全面地考虑你的问题,而且比一般人发现问题的眼光更犀利。所有权会提高责任感,如果你和导师分享研究成果的所有权,你肯定能得到他更多的时间和更好的指导与反馈。

最后,与一个在某一领域有经验的导师合作是特别有好处的。进行实证研究的教师一般都有实验室、仪器和一些科研经费(例如付给被试的费用)以供支配。许多人会允许你使用他们的部分或全部资源。另外,如果你的导师在某个领域中比较有名气,那你也更容易得到一些工作和实习机会——如果你做得好并得到他/她的强力推荐的话,你会获得更大的收获。

当然,在导师的研究领域内工作也存在缺点。一些导师可能控制性太强,不考虑你的观点或是水平,要求你按照他们的方式去完成工作。你会发现自己只是在服从命令,而不能锻炼独立的研究能力。即使硕士论文比博士论文少一些独创性,你也应该学习些研究技能而不是只成为一个"机器人"。如果今后想从事研究工作的话,你必须提高自己独立研究的技能。另外,导师对于熟悉领域的研究工作的要求会更高一些。以上这些问题并不一定会出现,而且我们相信同一个熟悉相关领域的教师合作是利大于弊的。

如果你不想只同一名专家合作,那么可以考虑一下吸收多个相关领域的老师来参加你的学位论文答辩委员会。在学院中、学术交流活动中,以及周边的学院、社区里寻找这些人,与这些人在其各自研究领域的交流可能使你获得导师没有提供的知识。

评估潜在导师的技能

主持一个学位论文答辩委员会需要特定的技能,这些技术是为了指导研究进程、监督论文的质量。这意味着导师要审阅你的开题报告和其他报告的草稿,对需要改进的地方给出反馈,在你遇到困难时帮你解决问题。导师还要主持你的答辩委员会,平时要监督开题报告会和口头答辩会。一个合格的导师会有清楚的预期,提供具体的反馈(积极和消极两方面),在合适的时候批阅你的草稿,和你讨论工作计划的进度,在答辩委员会成员间存在冲突时进行协调。

预期和反馈

在大多数的学院中,一些老师的要求标准可能比较高,另一些则不那么苛刻,他们对论文是否需要重写和研究方法能推广到哪些群体等方面的要求也较低。然而,如果由一个要求宽松的导师来指导的话,你在完成论文上花费的时间恐怕会更多。

我们必须承认对自己学生的要求很严格。因为我们相信严格的标准会对学生们产生积极的影响。首先,因为你和导师已经考虑了研究中可能存在的主要问题并且做出了适当的选择,同时上交给导师一份符合其严格标准的文章,那么你也不容

易因为论文质量不过关而在委员会其他成员那里出现严重问题。其次,因为你和导师已经就委员会其他成员可能提出的问题讨论过很多次了,所以你也为开题和答辩做好了充分的准备。最后,如果委员会其他成员对你的论文扔出一颗"炸弹",对你的工作感到满意的导师可能会帮忙解决这个问题。

除了考虑教师的要求标准之外,还要考虑一下未来的导师可能提供的具体建议和反馈的内容。因为在完成毕业论文过程中最令人沮丧的事情之一是有人告诉你文章和研究需要改进却又不指出究竟要在哪个方面做出改进。你不应该指望老师给你找到具体的参考文献,帮你做出所有的决定或者是重写论文中那些糟糕的句子。但是你应该希望潜在的导师清楚地告诉你在完成论文的整个过程中哪里是需要改进的地方。

时间表和有效性

在导师同意学生提交论文到答辩委员会之前,大多数学生都对学位论文中需要改进的问题得到了不少的反馈意见。得到这些反馈的时间是非常重要的:几个月来执著于修改草稿中引言和方法部分问题的导师会耽误你大量的时间。导师应该在恰当的时间和你见面并且讨论论文中出现的问题。如果一个学生需要讨论问题或者需要预约一次会面,我们一般安排在 1 周到 10 天之内完成这次预约。

一些事情可能会妨碍导师及时给你提供反馈意见,就像你考虑挑选未来的导师那样仔细地考虑这些事情。会有很多学生来和那些受欢迎的导师们安排预约见面,一些老师可能还会忙于大量的研究和交流活动,或者因为忙于咨询工作和指导临床实习等而无法抽出大量的时间给你。而且许多教师和学校一年只有 9 个月的工作合同,因此暑假期间他们可能不在。

如果你担心这些问题,告诉可能成为你导师的人;问问他们一般给出反馈的周期是多少,他们的日程安排是否能够保证经常和你见面;如果你们一起工作的话,你想多长时间见他们一次,等等。本章在后面的部分就如何从潜在的导师那里得到信息来重点讨论这些问题。

自信的技巧

我们希望你已经选择了能与之很好合作的导师和答辩委员会成员,但是如果还没有的话,你该怎么办呢? 如果一个委员会成员提出了一个不合理的建议,你的导师有足够的自信来支持你的决定吗? 如果进行研究时你在寻找被试方面出现了问题,你的导师会帮你解决吗? 在研究遇到困难的时候导师的自信程度对你的影响非常重要。此外,你要获得科研管理机构的支持,并且安排和被试研究审查官员会面等工作。如果你的学院是一个培养人际冲突的"温床",学位论文答辩会就可能变成了教师之间的"射击比赛",而参加答辩的学生则变成了不幸的旁观者,那么找到一个自信的导师则更为重要。最后,记住你的研究内容,完成一篇学位论文既是对你作为一个研究者应具备的能力的检验又是学习和积累经验的好机会。总体来说,硕士论文更像是一次对科研经验的学习而不是检验,博士论文则恰恰相反。在你进行研究项目的时候老师是怎么教你的? 他是不是帮你把大部分事情都做了,你是需要在很少的指导下自己寻找解决方法呢,还是在一个技术娴熟的导师的指导下提高

自己的研究水平?

调查可能的导师人选

3个有用的资源提供了关于选择潜在导师的信息。第一个是你的个人经验:你可能上过某个老师的课程或是与其交谈过,最好是曾经和他们一起工作过。这是获知在其指导下能否很好地工作的最好方式。

如果你在研究生生涯早期阶段读过本书,那么回想一下与哪些有共同兴趣的老师一起工作过,这样你可以获得第一手的参考资料。成为某些老师的有酬实验助手或是无酬志愿者也可以获得这些信息。如果当过无酬志愿者,在选择导师之前,一定要保证已经对老师们进行了准确的了解。如果你是为了了解特定的老师而投入其手下工作,你需要与其本人联系——而不只是忙于为他实验室中的博士后们当无偿的研究助理。

认真思考之前你和潜在的导师进行的互动:你做了什么,什么是你不喜欢的经验? 注意,你对导师的要求应该与对其他老师不同。例如,你可能喜欢史密斯博士的讲课风格,但是史密斯博士可不会在你的学位论文答辩会上发表演讲。要想想史密斯博士是否曾经给你的文章提出了具体有用的反馈。当然,史密斯博士可能不是在所有情况下的表现都一致,所以你需要找到其他信息来支持你的决定。

第二个资源是从和这些教师合作过的同学那里获取有用的信息。这里我们不讨论那些流传的小道消息。小道消息,就像杂货店出售的小报一样,一般是些娱乐信息,这些信息不可能完全准确。因为小道消息几乎全是由学生们杜撰或流传的,所以具有片面性。和生活中的大部分情况一样,在教师—学生事件中,每一个问题都有两面性,真相就在这两面之间。但是那些抵制学生流言的教师也从来看不到这些流言可能反映的一些事实。

要获得更多的可靠信息,就要和那些得到过某些老师指导的学生们谈一谈。毕竟,他们有这些"导师—学生"关系的直接经验。让这些学生说说他们的经验,问问他们你感兴趣的问题:论文每一个部分写了多少遍草稿? 老师多长时间给学生提供反馈? 老师是否及时接听电话和回复邮件? 老师们都给出什么样的反馈? 答辩委员会都提些什么问题,导师的反应是什么样的? 注意这些问题是关于教师的具体行为而不是学生的主观印象。尽管知道学生们喜欢或是讨厌导师的原因可能会很有趣,但是更重要的是知道导师和学生是如何进行互动的。无论你的"线人"是谁,知道了这些信息,你就可以决定是否与某一名老师进行合作了。

另外要结合自身的特点来选择导师。弱一点的学生会认为理智严谨的教授太"挑剔",而强一点的学生认为和这样的导师一起工作很刺激而且对自己更有帮助。对任何教师来说要完全满足学生多样的要求是很难做到的。因此请记住,导师们也不是随便地指导学生的:他们的选择偏好也在起作用。从我们的个人经验来讲,强一些的学生倾向选择那些对论文要求比较多但是更公正的人担任导师;弱一些和懒一点的学生经常选择可以允许他们走捷径的老师。但是没有人愿意与那些不负责任、专断、经常改变主意的人合作。

最后的信息来源是亲自与教师们会面。在某一时刻,你可能会准备一份导师候选人名单以及要问的问题列表。一名教师如何回答你的问题也给你提供了有用的

信息。与教师接触是学位论文礼仪中的一部分,也是你现在的主题。

与潜在的导师接触:学位论文礼仪 I

你已经考虑过想从导师那里得到什么帮助了吗?你和其他同学谈过了吗?你已经至少有了一个可行的选题领域了吗?现在是选择一个导师的时候了。随着你在完成论文的“战斗”中不断地突围,老师们也在选择那些他们愿意担任导师的项目。即使你已经在某人的实验室工作了,你也不能想当然地认为这个老师就会自动地成为你的导师,那只是你的预期罢了。你需要向他们询问一下,大多数导师关于选择指导哪些项目都有着明确的或者潜在的标准。例如,我们的一个同事提出了选择潜在学生的标准:

1. 研究设计必须包含具有比较作用的内容。
2. 这些比较内容必须包含有关文献中产生的假设。
3. 除了自我报告之外还必须采用其他的测量方式。
4. 必须至少上过一门我讲授的课程。
5. 学生必须具有必要的方法学知识以完成其个人的研究。
6. 学生必须有充足的时间来:①持续地为学位论文进行工作;②在教师的其他办公时间会面并讨论工作进度。
7. 在项目完成前,学生不得长期离开此地。

不管你是因为想获得更多的信息来与潜在的导师进行接触,还是要确定导师人选。从两个方面来说,第一步都是与老师预约然后请其指导你的论文和主持答辩会。在这时候应该直接把你的日程表交给他/她。如果你还在观望,你也可以向老师们这样表达你的想法。

> 我在考虑学位论文选题,其中一个领域是________。我想和您谈谈,看您是否有兴趣在这个领域中对我进行指导?我也期待能得到您的其他指导。您有时间和我讨论一下这个问题吗?

如果你想要这个老师来指导你的项目,可以这样说:

> 我想在________方面做学位论文。我想就具体的想法和您讨论一下,不知您是否有兴趣指导我的项目?

对老师可能当场向你提出的问题或做出的决定有个准备。例如,我们通常问学生,我们并不熟悉他们的主题,他们为什么会对我们(而不是其他老师)的指导感兴趣?这可以让学生们知道比起其他老师来,我们没有什么经验,然后节省双方的时间。同样,如果我们对一个主题感兴趣或者我们熟悉这些学生并愿意与其合作的话,也会立刻作出答复。如果我们的日程已经排满,或者对这个主题不感兴趣而不想与之合作,我们也会马上告诉他们。如果你已经深入地考虑过这些问题,那么询问潜在导师是否愿意阅读一下你写的关于研究设想和研究方法的简略概要(2~3页),这可以让老师知道你的写作水平如何。要确保你写出的东西清晰精炼,使用正确的术语,避免排版和其他错误。

如果教师同意和你见面,就带上日程表去赴约。如果你在观望,就带上一些能

帮助你决定他/她是不是合适做导师的问题去。可以从简短地介绍你感兴趣的领域开始,找出教师隐含的和明确的指导标准:他们一般如何与学生合作,在完善选题上投入多少精力,他们和什么样的学生合作得最好,他们现在是否有一些事情急于处理。除非你有一个具体的研究问题,不然不要期望有老师会答应指导你的项目。如果他们认为你的研究不重要或是对此不感兴趣的时候,一般也不会愿意投入时间和精力来指导你。如果你缺乏好的想法,问问老师们有什么建议或者有哪些读物可以帮助你勾画出一个研究问题来。如果有老师同意和你见面并进一步讨论和完善你的论题,这将是很好的进展。然而这并不代表着他/她就会指导你的项目。

如果有教师对指导你的项目表示出潜在的或者绝对的兴趣,并且你也有意请其指导的话,用"接下来我要做什么?"这样一个问题来结束会面。这给了教师一个机会来预期你需要做什么以保持论文的工作进度。如果你还在选题和观望,你可以告诉老师自己还在考虑论题并且要和其他老师谈谈,如果你想进一步讨论的话会再次来找他/她,接下来感谢他/她能和你会面。

要与导师确立清楚的规则、预期和直接的联系,对这些无论怎样强调都不为过。清楚你想要的是什么,能使整个会面过程变得不那么模糊。教师们可能不会提供准确的信息,所以学生们需要考虑如何恰当、直接地就项目计划包含的问题与教师沟通。你和导师可以在早期对研究前景和规则等问题进行一个交流,这可以为良好的工作关系做铺垫,并且也能防止由于导师和学生交流中的误解造成的问题。

预期的麻烦

在寻找导师的时候一些问题可能会引起麻烦。下面,我们列出学生一般要面对的 3 个问题,以及我们的处理方法。

1. *"我在麦克积博士的实验室工作,麦克积博士指导了我的硕士论文,但是我现在想让琼斯博士指导我的博士学位论文,有什么方法让我做到以上这一点但是又不疏远麦克积博士吗?"*许多学生担心如果离开某人的实验室就会放弃与此教师的良好关系,甚至不能从他们那里得到推荐信。其实这些担心是不必要的。在许多学校里,学生们在研究兴趣转变或是想接触另一个研究领域的时候是可以更换导师的(在一些地方甚至还会得到鼓励)。尽管教师可能为失去培养数年的学生而感到惋惜,但绝大多数都会对此表示理解。

在这种情况下,我们建议你和老师就这个问题谈一谈,下面的一些话可能比较有用。

> *我在考虑我的学位论文主题。尽管我从您这里学到了很多东西,但是我真的对________更有兴趣。我计划和琼斯博士就我们能否一起工作的问题谈一谈,但是我更希望先和您说明一下情况并且想知道您对此的看法。*

和蔼的教师会告诉你:"我很想留下你,但我尊重你的选择",而不会导致什么负面后果。实际上,老师都比较喜欢你在和其他人接触前先和他们交换一下意见。另一些教师可能会告诉你他们花费了很多时间来培训你,他们要求的"回报"就是你做一个以他们感兴趣的主题为内容的学位论文。然后你再决定是否为了保持这种良好的关系而更换你的论题和研究计划。

2.“我想研究太阳灯对虫子转弯行为的影响，巴顿博士是那个领域的专家，但是我不想和他一起工作。”这里至少有两个问题：首先，除非有人分享了巴顿博士的研究经验，否则你不会找到比巴顿博士更合适的人来指导你的论文。第二，巴顿博士很容易注意到你的研究主题并且发现你基于某种原因选择另外一个人做导师，尽管巴顿博士本人对此是否有意见并不确定，但是，和巴顿关系密切的教师可能不会愿意担任导师或是答辩委员会成员。另外，人们会好奇为什么你没有选择这个领域的专家作为自己的导师。你觉得自己有什么好的说辞来解释这种行为吗？

我们建议在解决这些争议的时候考虑两个问题。第一，确定不选择巴顿博士的原因。有什么方法可以将问题最小化吗？在你选择导师的时候可以讨论一下细节的问题：“我很想在您的领域中工作，但是您看起来很忙，而我可能需要很多指导。如果我们一起工作的话，可以安排定期的会面吗？”论文的逻辑问题可以通过充分的开放性讨论来解决，然而不负责的行为和负面的人际风格就没法通过深入的讨论来预防和解决了。第二，对选择其他人而不是巴顿博士所产生的后果进行反复的、现实的评估。巴顿博士不会介意吗？巴顿博士会如何反应？巴顿博士的反应或看法对你毕业之后的生活有什么重大影响吗？许多学生会把没有选择巴顿博士的影响灾难化，但可能巴顿博士很高兴不用指导一个不喜欢的学生！同时，如果你想在巴顿博士的领域里找到一份学术性工作而他已经是一个国内知名的专家，那么明智点还是选他做你的导师吧。

或者，你可以选择其他研究题目。你当然能够再找到一个感兴趣的题目，这样在选择其他人做导师的时候你和巴顿博士都可以保留一定的余地。在完成学位论文之后你可以随便更改自己的研究兴趣。

3.“我找不到任何人来指导我的项目。”如果是这样的话，那么多分析自身原因。以我们的经验来看，造成这种情况的原因一般是选题不当。教师会因为你的论题不重要、太差或者不在他们熟悉的领域而拒绝指导你。如果学生一心扑在一个特定的研究问题上，在没有教师的指导下进行工作，然后再把这个“既成事实”提交给老师的话，他们一般会拒绝。解决的办法就是尽早地和教师进行接触，谈谈你们两个都感兴趣的题目。然后让老师知道你的想法很灵活，而且对于哪些东西重要、哪些不那么重要保持一个开放的态度。

另一个找不到导师的原因是学生没有进行充分的调查，一门心思想让某位教师指导你，结果却发现他/她不能再多指导一个学生了，要去度假或是因为其他一些原因不能指导你。最后你只能得出一个“我找不到导师”的结论。你可能需要拓展自己的论题以增加别人与你一起工作的兴趣。

最后，一些学生找不到导师是因为学术或个人问题。那些在训练中反复出错的学生可能发现很少有教师愿意与他们一起工作。如果你也有这样的困难，那么应该诚实地看待这个情况。首先，需要找到出现这些问题的原因。是你不够可靠、不友善、性格上棱角太多还是不够变通？你是否因为缺少研究生水平的基本学术技能而需要进行额外的补习？第二，你需要知道在短时间内是否能够改变你的技术水平、行为、态度等。是否需要参加高级咨询、治疗或是寻求其他教授的协助。如果你能达成这两个目标，那么不断地与导师进行交流是改变现状最有效的方法，向老师展示你至少已经在某一方面取得了进步，然后再要求得到一次机会。要改变负面的名

声是很难的,你需要更加努力并且比其他学生做出更多的妥协,以此来逐步改善老师们对你的不良印象,并请他们来指导你的学位论文。

如果你认为自己不能在读书期间做出成功的改变,考虑下面两个建议:①休一个假来解决问题;②退学,然后找一份合适的工作。尽管后者是一个极端的方法,但是在研究生院里的错误同样也可以预示着在未来工作中犯的错误。如果你并不适合目前的工作,那还值得待在研究生院吗?或许找到能发挥你特长而不是暴露缺点的事情来做会更好!

完善研究问题

我们假设你已经成功地确定了一个研究方向(例如牙医恐惧症、儿童性侵犯、抑郁症、HIV 预防等),并且还能得到一个颇有名望的教师的指导。接下来你要做的就是将研究问题的范围缩小到可操作的程度。要做到这一点,你应该先在这个领域中构建某些可研究的问题。这些问题的**形式**需要非常具体。我们在后面会对此作出解释和说明。首先,我们从这些研究问题的**来源**开始本节。

一般来说,你提出的问题也应该是自己所在的科研团体感兴趣的问题。仅仅提出一个详细的研究计划并不见得就好,因为这个主题可能缺乏研究价值,还没有人在此类问题上发表任何东西也许就是这个原因。更重要的是这不足以作为这一问题应该被提出或是得到解答的首要原因。

就一个可研究的问题而言,它需要来源于文献。也就是说,你需要知道它的发展、演变、概念和方法论等,要知道科学是一个知识积累的过程。对于这个问题的回答会让我们得到哪些认识呢?它为什么是重要的呢?诚然,鉴别出学校里存在与学习障碍相关的认知缺陷的学生比例是很重要的。然而除了对于那所特定学校而言,这一研究结果并没有什么意义,别人对于其他相关问题的兴趣可能要远高于此。如果你真的想充实自己的研究,应该选择那些对具体应用有理论意义的变量,如学习障碍的特点等。记住研究的目的是找出普遍性的知识。如果研究能建立在先前的研究和理论之上是最好的。我们在后面的章节中讨论研究假设形式的时候会加以详细的说明。

不良研究思路的来源

在我们开始讨论具体研究构想的好的来源之前,给大家一个忠告:研究思路一般产生于个体的研究经验,如酗酒的人在康复之后决定研究酗酒者的家史;有慢性考试焦虑的学生选择考察数学考试时考试焦虑和试卷的问题类型的关系;已经 40 岁的成年研究生决定研究商务主管配偶的性衰退现象等。事实上,我们中许多人之所以从事行为科学研究,特别是心理学领域的研究,都是因为对我们自己的经历感兴趣。不要误解我们的意思,首先这是一个进入心理学领域的好理由。但是因较高的个人情感卷入而设计一个学位论文研究可能就不是个好想法,还是把这些问题留到心理治疗时再说吧。每次当你提笔写作时或是打开一个被试的数据时,要避免自己进行大量的个人反思是非常难的。

除了妨碍你的进程以外,个人的问题是不大可能通过独立的、客观的和必要的

分析观点应用到科学中的。你可能会站在被试的角度上看待一些问题，这在某些程度上会干扰你完成研究工作。最重要的影响是出现“自我兑现预期”的现象（Ambady & Rosenthal,1992）。如果你从个人经验出发说服自己——焦虑症学生在考试时在多项选择问题上的表现比单项选择问题上差，那你可能不自觉地设计出能得出这一结果的研究。有研究显示，实验助手可能会因为实验者对他们的要求而影响实验数据的偏向（Kent, O' Leary, Dianment & Dietz, 1974; Shuller & McNamara, 1976）。当然，我们可不是建议你故意这么做，只是告诉你做研究时很容易出现这种问题。

选择的研究主题如果引起了强烈的情感反应也会以某种方式影响你的客观判断。例如，在委员们的建议和你对问题的个人理解相冲突时，即使非常好的建议也很难被接受。这些会致使你表现得死板、抵触和不灵活，而教师们却不知道真正的原因所在。

良好研究思路的来源

你可能从我们之前的评论中猜出来了，最好的研究思路其实来源于你已经做过的研究，因为你既有这一领域的知识又有研究的经验。再者，如果你与某一个教师有过研究合作，那么他/她给你提供的建议可能发展成为很好的主题。最后，前人的研究总是能激发你对很多问题的思考，其中的一个很可能就成为了学位论文开题的基础。

前人的文献，特别是近期的文献，是很好的问题来源，多看一看。如果其中一篇激发了你的灵感，花一些时间考虑一下如何将灵感发展成为你的研究对象。与你尊敬的老师或潜在的导师候选人讨论一下这些想法，观察他们的反应。如果大部分的反应是积极的，那么在你已经全面研究过这个主题之后，与有关文章的作者联系，看看他们在这个领域正在从事哪些工作。详细地打听出你希望做的关于那个问题的细节，问问作者是否已经做过此类研究或者他们是否知道谁正在做。如果对这两个问题的答案都是否定的，进行一点探究看看作者是否知道问题的答案。此刻这个研究问题对你来说是多么好的一个想法呀！

这点儿调查工作是非常值得做的，而其可以节省大量的时间并避免陷入可能的窘境。比如，原来的研究者因为同行指出了逻辑上的内在缺陷而放弃了研究；或在专业报告会上，被一些听众严重质疑了研究的道德问题等；或者你可能听说一些人已经投稿发表他们的研究结果了。这样的消息难道不值得发封邮件或是打个电话过去询问一下吗？

研究问题的另一个很好的来源是把某一群体中的研究范式应用于另一群体。例如，Wolfe, Gentile, Michienzi, Sas 和 Wolfe（1991）使用了事件冲击量表来评估对创伤性事件的具体反映，此量表（*Impact of Events Scale*, Horowitz, Wilner & Alvarez, 1979）是后来修订的儿童事件冲击量表的基础。对行政型妇女领导风格的研究可能很容易产生出其他的相关研究，例如可以更换一下研究群体，因为行政人员是一个不具有代表性的群体。

文献综述是另一个找出自己研究问题的途径。《心理学公报》①就发表了很多

① Psychological Bulletin. ——译者注

这样的文章,《心理学年度评论》①,《临床心理学评论》②和《发展评论》③是这些综述类文章的另一来源。许多书籍也包含了文献综述。这些文章的作者经常发表对一些问题的深刻理解或是指出现存研究的不足,你为什么不进行一个能补救前人研究缺陷的研究呢?

在另一些情况下,验证已经发表的研究可以为接下来的研究提供一条康庄大道。我们在早些时候提到过,验证前人研究的硕士论文比博士论文较能让人接受。推荐你"复制"前人的研究有很多的原因。如果最初的研究是个经典研究,我们会从中学到很多东西。实际上,唯一值得重复的研究是那些经典的实验研究,因为重复时不会犯前人的错误。查查硕士学位论文规范,看是否接受重复性的研究。

另外,你也许会发现一些文献之间的联系。很多其他专业的人士和心理学家研究相同的问题,其中包括社会科学家、人际交流领域的科学家、物理学家等。即使是在心理学中也有数不清的未整合的学科方向。把分离却相关联的东西整合到一起是一项挑战,而且经常能得出一些新奇而有价值的东西。

最后一个研究想法的来源是其他人的学位论文。我们知道一些教师会故意引导学生不要考虑之前的那些学位论文。发生这种情况的原因是教师们不确定由其他人指导的研究质量究竟如何;另一部分原因是这些学位论文没有经过像绝大部分已发表的文章那样严格地评审,因此不能够拿出来作为研究的模板使用。教师们认为,如果一个研究足够好的话总会被发表,学生可以在期刊文献中查到。

我们的观点是已经存在的学位论文**可以**给未来的研究者提供很好的信息来源。我们通常建议学生们查阅几篇这样的学位论文或是开题报告。这不但可以让学生对自己的想法做一次现实性检验,还可以帮助学生找到可以作为模板进行参考的目标。如果学生们提出的是脱离前人学位论文研究的扩展性问题,我们也建议把已经完成的学位论文作为他们的参考。而且像在第1章中提到的,在潜在导师的指导下完成的学位论文可以作为对你提交的成品质量是否符合他们要求的参考。但是我们不建议把图书馆架子上的学位论文或是国际学位论文摘要当成研究问题的主要**来源**,浏览那些已经发表的文章可能产生更好的研究想法。

将研究问题转化为可研究的形式

一旦决定了主题领域和要研究的问题,你就要用语言很好地表述这个问题并使其适合研究。表达良好的问题具有共同的特征(Kerlinger,2000)。要记住最简单的表述标准就是:第一,以问题的形式叙述你要研究的内容。第二,需要提出假设检验的关系,这要求在问题里必须明确需要检验的主要变量。因为研究的目的是为了推动科学的发展,而科学研究的是变量间的关系,如果你不是在检验一种关系,就不是在从事科学工作,至少在本书的定义下是这样。没有科学,就没有学位论文,就这么简单。

假如你的研究领域包括儿童性侵犯,你希望研究儿童对具体事件的记忆容量。你的问题可能是,"未曾受到性侵犯的儿童在玩带有性别特征的娃娃玩偶时会更多

① Annual Review of Psychology.——译者注
② Clinical Psychology Review.——译者注
③ Development Review.——译者注

地描绘出性行为吗?"这就是一个问题的形式,它是否反映了一个需要检验的关系?自变量就是娃娃玩偶是否具有第二性征。把儿童暴露于两种不同性别特征娃娃玩偶的条件之下的行为作为性行为的参考来检验,儿童的行为构成了因变量。研究问题提出的是一种自变量和因变量的关系。因此,它符合我们的第二个标准。

第三条标准要求研究问题具有实证检验的可能性(Kerlinger,2000)。如果问题符合第二条标准,即,提出了一种关系,它就有一半符合了第三条标准。然而,光提出一种关系是不够的,这种关系还要可以进行实证检验。可检验的理论基础是研究中的特定变量要与研究问题相关。这些变量有操作定义吗?如果变量不能被操纵和测量,研究中的问题就不可能得到解答。如果它们可以被操纵,那么问题提出的那些关系就可以得到实证检验。

你将要花费一些时间来提炼和改写研究问题以使其足够清楚和具体,并能保证研究的继续。这个过程会迫使你全面地考虑自己究竟想要知道些什么东西。你可能从一个模糊的问题开始,"家长和青少年如何就性问题进行交流?"这个问题并没有暗示出任何关联。即,没有出现自变量和因变量。可以把它改成:"家长一青少年如何就涉及青少年冲动型的性行为的问题进行交流。"如果你确定了研究家长和孩子关于性行为问题进行交流的具体方向,可以再进一步修改。

现在你有了一个研究问题,接下来让我们花几分钟谈谈被试的可用性。被试的选取是如此重要,以至于我的同事会建议学生们先找到被试然后再决定要做什么研究。如果被试不可用,最明智的决定就是不要去做这样的研究。你在工作中(或者实习中)接触的人符合你的研究需要吗?你的父母、同龄人或者朋友和这个研究有联系吗?你可以从潜在导师那里联系到被试吗?本书的一个作者指出,处于一个系统外部的人经常缺少收集系统内部数据的能力,进而影响研究的现实性。所以一定要保证你能找到需要的被试。如果研究问题要求选择特殊群体(例如,癌症患者、流浪者等),能找到这些被试就更为重要了。但是在使用相对容易获得的群体时,这个问题就不那么重要了。

做出措辞严谨的假设

符合3个准则且措辞良好的问题可以帮助你把假设具体化。假设是用来推测两个或多个变量间关系的声明性句子(Kerlinger,2000)。良好的假设可以从研究问题中直接获得。例如,我们前一个问题:"未曾受到性侵犯的儿童在玩带有性别特征的娃娃玩偶时会更多地描绘出性行为吗?"就这个问题,我们的假设可能变成这样:"未曾受到性侵犯的儿童在玩带有性别特征的娃娃玩偶时会更少地描绘出性行为吗?"注意在两个例子中,我们的假设开始于基本的研究问题,然后发展成由问题确定的变量间关系特性的具体预期。也要注意假设可以推测目前存在的行为与原始状态之间的关系,因此,使用现在时组织句子。我们不假设它与原始状态可能存在的关系,所以要避免使用将来时。

此时,你可能会考虑刚刚提到的一些假设,这些假设可以顺利地得到研究结果。你可能认为这很简单,但事实却不完全是这样的。你需要说明预期是根据什么产生的,一般有两个来源:①前期的实证研究;②根据理论作出预期;或是两者兼有。如果前面的研究形成了特定的研究问题,这些研究将会延伸出更具体的假设关系,并进行

检验,最后使问题更加清晰。同样的,理论也会延伸出有待通过检验来确定的关系。

除此之外,还有一些研究领域过于前沿以至于假设关系很难被发展出来。另外的一些领域(如行为学或天文学)比起其他领域来说则拥有更长的研究历史。如果你在建立假设上有困难,而且熟知文献以及相关的理论框架,那么就此问题和导师谈一谈,千万不要虚构假设关系。科学包括独创性,但不能产生毫无根据的假设。

另外,研究假设一般产生于文献阅读之后。你可能认为在读完了一些文章之后可以预期自己的研究结果,但也可能发现随着主题越深入,这些文献中矛盾的地方也越多。这是很正常的,是科学进程的一部分。进一步阅读相关文献,思考自己的研究主题,提炼、修改你的假设。

我们还需要了解假设的类型。我们经常会遇到"虚无假设"和"研究假设"这两种类型的研究假设(Ray,2006)。研究假设决定了研究中需要确定的关系,而由罗纳德·费舍爵士(Sir Ronald Fisher)最先提出的虚无假设是假定变量之间不存在关联。尽管你对自己的某些研究假设十分感兴趣,但是统计方法却不能验证它们。统计方法只能对虚无假设进行证伪。多数导师都希望学生在开题报告中加入关于研究假设的内容,但并不要求他们把虚无假设陈述出来。

注意,把一个虚无假设作为你研究的预期结果一般是不能够被接受的。因此,"男孩和女孩在科学故事上的兴趣相同"这样的假设是不够合理的。因为统计学的作用是证伪虚无假设。寻找你期望存在的关联,而不是寻找某种关联的缺失。当然,你会发现在一种条件或是分组下产生的关联与另一种情况下不同,这个假设就是个好假设。因此,"不同性别的兴趣依赖于刺激类型,男孩对科学玩具的兴趣比女孩更大,但是男孩和女孩对科学故事表现出同样的兴趣"就是一个合理的假设。这是因为它对性别和刺激的种类做出了显著交互的预期,这是可以被推断性统计检验的关系。

谨慎地对需要检验具体关联的研究问题进行假设。假设也表明了要检验关系的性质,因此,"在教育水平和对自由职业的偏爱上存在关联"的表述就不如"在教育水平和对自由职业的偏爱上存在正相关"这样的表述合理。表述清晰的假设所需要具备的第三个特征也很重要:它应该包含研究设计的本质内容,在上面的例子中暗含了一种相关设计。另外,假设还需要包含研究群体的范围。"在职业女性中,教育水平和对自由职业的偏爱上存在正性关联"的假设就符合规范。

现在应该停止关于假设问题的进一步讨论。然而,对于研究新手来说再扩充一些内容是很正常的。其中一个普遍存在的错误是在假设中包含了对检验变量所采用的具体测量方法的描述。例如,"在职业女性中,教育水平(由霍勒斯·曼教育成就量表测得)和对自由职业的偏爱(经 ACLU 量表测得的,这种测量的思考过程十分清晰)存在正相关。"诚然这个假设是非常具体的,但是表述却还有一些笨拙。尽管一些人可能不同意,但是我们相信过犹不及的描述会使我们失去研究的要点。我们研究的主要目的是考察变量间直接的关系,并以此来完善一门科学。但是有时过分关注研究方法中的具体测量技术是不恰当的。只有在这些具体测量技术会影响主要变量时,我们才关注它们。就像其他人说的那样,"假设应该是指我们感兴趣的变量,或是具体的实证定义,而不是一些具体的指标"(Pedhazur & Schmelkin,1991,P. 179)。总而言之,要基于你的归纳研究发现的层次组织假设。

另一个需要避免的错误是在假设中加入具体的统计检验名称。因此,"预期单

尾方差分析将会揭示律师、心理学家和会计外在信仰程度的区别”是错误的表述方式。在这种情况下,我们表现得好像是对由统计方法引起的职业组外部信仰的差异感兴趣一样。我们应该通过统计来检验差异,而不是用它们来说明假设。

学生们经常询问,“我应该提出多少个假设?”我们的回答是:“不要太多。”有时我们看到开题报告中包含了 10 ~ 20 个假设。通常这些假设中包括对每一个刺激水平的预期和扩展。这么多的假设给学位申请者打上了一些非常清楚的烙印:①想要的东西太多了;②没有为了缩小具体研究假设而深入思考;③对总体把握的程度很低。在做学位论文时要牢记“奥卡姆剃刀原则”:保持最简,没有那么多独立的行为现象。进一步来讲,如果你需要检验 10 ~ 20 个假设,因素分析可能帮助你发现其中的关联,并且把范围缩小到 3 ~ 4 个独立因素之内。当同样的现象只能由极少的概念或变量解释的时候,科学便有了进步。应该经常反思:我们形成的假设是不是都在反映不同的事情。如果不是,在你要重点研究的问题上组合和修改它们。如果它们反映了太多不同的事物,那么你的研究目的就过于繁杂了。这时,你应该和导师一起重新查阅文献来控制一下假设的数量。

专栏 4.1 总结了我们推荐的一些关联假设,可以帮助你判断提出的假设是否合适。

现在,你对于研究假设能揭示出哪些关联的实验设计已经有了一些眉目。此时假设的性质会直接决定你该选择哪种统计检验方法。最后一个问题:你已经为要进行的研究做好充分的准备了吗? 例如,研究新手有时会进入一个他们没有决定好对研究的理论变量或结构采用何种方式进行统计检验的领域。这时我们需要考虑一些问题:①“你上过哪些量表结构和测量理论课? 做过有关因素分析的作业吗? 或者概括理论的水平如何?”②“计划用多长时间完成这个项目?”③“准备用自己制作的量表来完成学位论文吗?”我们的教训是,避免让研究的问题与你的实际技术水平脱节,一定要做力所能及的事。

专栏 4.1

假设清单

是	否	
□	□	1. 你的研究假设提出了两个以上变量之间的关联了吗?
□	□	2. 你的研究假设指出了具体的关联性质了吗?
□	□	3. 你的研究假设是用现在时表述的吗?
□	□	4. 你的研究假设揭示研究设计中的关联了吗?
□	□	5. 你的研究假设针对的是你要研究的群体吗?
□	□	6. 你的研究假设是在归纳已有发现的水平上组织的吗?(如指的是研究的结构而不是具体测量)
□	□	7. 你的研究假设反映了变量间的关系而不是它们的具体统计检验名称吗?
□	□	8. 你的研究假设对其他非必需的方法学细节适用吗?
□	□	9. 你控制了研究假设的数量了吗?(如 5 ~6 个或者更少)

注意,我们对每一个问题最满意的答案是“是”,除非你的导师或是院里的规定与此相反。

召集答辩委员会成员

如果你有一个具体的研究主题,有关于研究假设的思路,有导师的指导,那么接下来就是要确定学位论文答辩会的时间。学位论文答辩会的规模、构成和具体职能因学院的不同而不同。但是委员会成员的一般职能是一致的:①提供包括对文章写作框架在内的最后修改建议;②起到对开题和论文质量进行检查的作用;③在学生答辩时起主导作用。在硕士论文答辩中,委员会通常希望学生能理解并且能够明智地谈论研究问题。在博士论文答辩中,委员会将评估学生作为一个独立的研究者的能力。因此,答辩委员会起到顾问和质量检查的作用。不同学校和不同委员会对此的平衡不同,这取决于是硕士答辩还是博士答辩。

理想的委员会成员应该起到辅助导师工作、产生研究新视野和想法的作用。例如,本书的另外一个作者就只同意在关于被试群体、自变量或是主要调查的因变量方面帮助导师。委员会成员的理论观点和见解应该与学生的相适应,而且他们应该对学生的研究主题有一些兴趣。学生在和招募被试的代理人紧密合作的同时应该考虑聘请一位博士级的代理人来担当委员会成员,这个人通常充当学生在代理机构中的辩护人和联系人,并且提供关于代理规定和选择合适的代理资源的有用信息。如果你要借鉴别人的统计方法并且需要更多的指导,就在自己的答辩委员会中加入一位统计学专家。

委员会成员还应该和导师有一个良好的互动关系。另外,委员会成员应该是公正的,有指导性并且可靠的。尽管委员会成员表面上称赞学生,然后在答辩时暗中破坏是很少的,但是你也应该尽可能多地收集相关信息,防止这种事情发生。

在考虑选择具体的委员会成员之前,找到关于委员会成员在准备开题和提交文章时都有哪些规定(成文与不成文的规定)需要遵守。许多学院遵循导师主导模式,即委员会成员只有在导师许可的情况下才能读到开题报告和完整的文章。这些许可拓展了学生和导师在文章成形之前的工作,委员会成员期待着对项目的具体方向提供各自领域的建议(例如,具体的测量工具或是对数据的统计处理)成为导师的替代者。然而,他们通常不会(除非在委员会成员比主席在这个领域知道得更多的时候),由导师,而不是委员会,在研究假设、研究设计和测量的选择等细节上给学生提供帮助。当导师对开题报告满意的时候,委员会成员才可以发表建议,除非他们在更早的时候就因为自身的技术水平比较合适而被要求提供帮助来解决可能出现的问题。

导师主导模式并不是唯一的模式。在一些学校里,委员会成员承担的责任和导师相近,也阅读由学生提供给他们的开题草稿和最终的报告(我们把这个叫做委员会主导模式)。不过我们的经验是,在这种情况下委员们的职责经常含糊不清。我们应该让学生来决定委员会成员应该什么时候、如何参与到这个过程中来。当这个程序按规定必须遵守的时候,学生可以通过询问导师的意见来选择工作模式——导师希望委员会成员承担哪些责任和提供什么服务。责任具体化是非常重要的,什么时候让委员会成员读到那些内容?哪一版文章?是否应该有一个正式的开题报告?接下来学生应该就这些预期与潜在的委员会成员沟通。

当责任含糊的时候产生的另一个问题是,学生应该怎样处理不同委员会成员对

某一研究问题的不同意见。这会使学生奔忙于主席和委员会之间，有时甚至误解或忘记了每一个人都说过些什么。在这种情况下，我们建议进行一次团体交流：安排一次包括全体成员的会议，然后面对面地讨论问题。

了解潜在的答辩委员会成员

我们对学生如何寻找导师的建议也适用于寻找委员会成员：找到组织委员会的正式规则和那些可以成为委员的人；从同学那里了解不同教师作为委员会成员的经验；大范围搜集关于他们的信息。那些“小道消息”经常包含一些虚假信息。我们就曾因误解和猜测导致不能与一些同事很好地相处！选择那些更好的消息来源（如，与不同教师合作的优秀学生的经验）。另外，关于潜在委员会成员的人选可以询问你的导师。如果你的导师和他们工作过，他/她会知道这些人对你的研究项目会有多大的帮助。在邀请委员会成员之前有必要先和你的导师谈一谈。这是因为你的导师可能会有具体的建议，而且他与某些同事一起工作会比和其他一些人一起工作效果更好。

你的导师还会建议你应该在什么时候去接触委员会成员。一般来说，我们建议学生在确定了研究问题、研究方法和准备好开题及进行研究的合格的日程表之后再去组织答辩委员会。我们要知道具体的论题，同时知晓研究的时间进程也很重要，因为教师可能会缺席，可能会去度假，可能不愿意参加太多的答辩会，等等。暑假无薪的教师很多是不愿意在这个假期和学生见面或是出席开题报告会和答辩会的。替换一个委员会成员是很难的，所以，依照你的时间表来组织一个能够保证全部成员顺利出席的委员会吧。

接触潜在的委员会成员：学位论文礼节 II

与前面挑选导师的建议一样，我们建议你有策略地会见和邀请答辩委员会成员。和那些潜在的委员会成员联系并交谈，看他们能否加入你的答辩委员会。给他们提供一份对研究主题、研究方法以及研究进程时间表等方面的简略概述；问问他们是否可以看一下关于你的研究内容的简短的概括性材料；你也可以告诉潜在的委员会成员自己希望获得哪些帮助。如果你希望潜在的委员会成员在具体任务（如在统计学或者是研究方法）上帮助你，告诉他们你的这些要求（如，“我希望您可以在联系学校方面帮助我”）。问问他们对你的项目可能提出的要求。例如，一个委员会成员是统计学家，可能希望只有在你的研究主题确定下来后才和你讨论统计问题。与找导师一样，你要在充分地收集了信息，并且确保潜在的委员会成员与你的研究项目相适合之后再发出邀请。

补充资源

Darley, J. M., Roediger, H. L., & Zanna, M. P. (Eds.). (2003). *The compleat academic: A practical guide for the beginning social scientist* (2nd ed.). Washington, DC: American Psycholgical Association. [这本书的一章从教师的观点对教师—学生研究合作提出了看法。对那些打算进入学术职业领域的人来说是一本很好的书。]

Frick, R. W. (1996). The appropriate use of null hypothesis testing. *Psychological Methods, 1*, 379-390.

Oetting, E. R. (1986). Ten fatal mistakes in grant writing. *Professional Psychology: Research and practice, 17*, 570-573.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.

Sternberg, D. (1981). *How to complete and survive a doctoral dissertation*. New York: St. Martin's Press. [为选择答辩主席提供了更多很好的信息来源,包括避免选择那些可能存在性癖好的人。]

Webb, E. J. (1961). The choice of problem. *American Psychologist, 16*, 223-227.

思考与练习

寻找主题和委员会成员

发展你的主题

☐ 选择一个一般的主题领域

☐ 和老师一起工作来发展你的想法

☐ 发展研究问题

—思考你参与过的研究

—避免超过个人承受能力的主题

—使用比较新的文献

—谨慎地使用其他学位论文

☐ 将研究问题变成可研究的形式

—组织问题

—确保问题提出了一种可被检验的关系

—确定问题是可被实证检验的

☐ 使用表 4.1 很好地组织假设的语言

集合你的队伍

☐ 明确导师和委员会成员的正式规则

☐ 明确导师和委员会成员的非正式规则

☐ 明确你想从导师—学生合作中获得什么

☐ 明确你感兴趣的领域中潜在的导师的经验

☐ 评价潜在的答辩委员主席的技能

—预期

—反馈

—时间表和可用性

—自信

☐ 了解潜在的导师

—考虑和这些人工作的经验

—和同学谈一谈

—和老师们谈一谈

☐ 接触潜在的导师

—明确和他们讨论的重要问题

思考与练习

—清楚地交流

—获得主席的承诺

☐ 了解潜在的答辩委员会成员

—同导师交谈

—和同学谈谈

☐ 接触潜在的答辩委员会成员

——提供你的研究概述和时间表

——获得答辩委员会成员的承诺

5

形成和交流研究计划：开题

此时，你已经知道自己想做什么并且也知道选择谁来指导你。接下来，就需要使自己的计划更加完善。在本章中，我们将勾画出开题报告内容和结构的轮廓。关于开题报告具体内容的解释包含在第6到第9章中。在第13章则讨论如何提交论文和进行答辩。首先，让我们来看看学位论文的开题需要完成哪些东西。

理解学位论文开题的作用

开题报告是关于你期望做什么研究以及如何做的文章,不论从哪个方面来说,开题都是很重要的。首先,开题报告是在进行学位论文工作的初期就需要完成的一个重要任务。从某种意义上来说,它算是后续研究的预备性练习。撰写开题报告可以为复杂的研究过程和选定的研究问题提供第一手资料。开题报告也给导师和答辩委员会成员提供了你为研究所做的准备以及自己是否具备成为一名合格学者的资格。因为开题报告奠定了整个学位论文的"基调",因此你需要如何处理开题这个交流过程以给委员会成员留下一个好印象呢?你想被看成是在研究的每一轮进程中都需要指导和安慰的无用之人吗?你是否希望作为一个有良好判断力的人而被大家所信赖?

论文开题是如此重要的另一个原因是,一旦开题报告被答辩委员会接受,它就将成为你进行后续研究的蓝图。开题报告确定了你应该如何完成研究计划的具体步骤。假如委员会赞同你的开题报告,那就是说,如果你确实按照开题报告说的那样完成了研究,你的学位论文就可以获得通过。

在一些学校里,开题报告被认为是一种非正式的合同(见 Sternberg,1981):如果你按照上面的流程进行研究,答辩委员会将不能在事后要求你增加实验被试、更改实验流程或者加入另一个控制条件等。实际上一些学校要求委员会成员在开题报告会上提出改进意见后再签署一份协议,并与修改过的开题报告一起放在院长或是其他管理人员的办公室存档。这样,数年后,当疲惫的博士申请人带着她完成的学位论文从斐济返回却发现教师中已经没有一个是她原来的答辩委员会成员时,这份存档的开题报告就是她所在的大学曾经批准她在南海岛屿对80岁老人进行水下有氧运动情况进行研究的直接证据,而不会因为缺少了当事人导致她的论文不被承认。

然而,不是所有的教师和学院都有这种合同。即使是那些有此传统的学校也会指出,对开题报告的许可不代表着学位论文就可以通过。你必须还要认真地进行研究,恰当地分析数据,提交出可被接受的研究报告,等等。不过大部分的研究者在实际进行研究时都会对开题报告的内容做出一些改变。尽管如此,对学生和老师来讲,开题报告也会使他们更加明确本校的一些有关规定。

你产出的作品就是展示自己是否有资格成为一名学者的证据。注意,尽管新颖性和独创性是评价理论和研究的一个重要指标,但在准备开题报告时特立独行的态度和行为却是不受欢迎的。开题报告会上没有足够的时间来给你挑战学院的固有规则和展示"变革"的勇气,还是老实地遵循学校对开题报告的要求进行吧。把你的创造性用在其他更重要的挑战上,然后按规矩办事。

开题报告对许多学生来说是一项可怕的难题。可能你自己都不敢相信,其实自己早就为此做好了充分的准备。你现在要做的就是整理所有的笔记和得出的结论,把为研究而收集到的所有量表归类成一个单一的文档。在下一部分里,我们将给出一个关于开题报告中应该包含哪些内容的概述。

了解学位论文开题报告的内容

一个开题报告里应该包含哪些内容呢?学院和教师对开题报告的要求详略不同。极端一点来说,学生需要提供一份包含少量研究细节和附录,以及简短的文献回顾和研究方法等内容的文章。宽泛点来讲的话,开题报告本质上是你将要提交的学位论文中前一半内容的草稿,其中还应该包括所有附录中的有关补充材料。如果你还没有开始准备开题报告,请导师给你推荐一些别人的开题报告作模板。把这些开题报告与最终的学位论文进行一下比较,可以确定出申请者的开题报告应该具有的详细程度。

此外,在你的开题报告中多加入些细节可以让导师给出更多的关于具体研究想法和计划的反馈意见。现在完成的越多,在答辩会上委员会成员对你提出的疑问就越少。例如导师可能会建议你重新挑选一些被试进行实验来修正那些可能出现的错误。

尽管导师和委员会成员对开题报告的要求有很大不同,但是一个好的开题报告需包括的 3 个部分是不会改变的:前言(文献综述),方法和结果。有些开题报告会把数据分析的内容放在方法这部分内容中,而不是安排在结果部分里报告。当然,有些内容有时会出现在标题、表格、参考书目或附录等不同的地方。尽管如此,以上 3 个部分仍然是开题报告的核心内容。下面我们来对开题报告的每一部分进行概述。后续的章节(本书第 6—9 章)会更加具体地讲到开题报告的这些部分。

前言(文献综述)

在前言部分,要谈一谈你研究的领域,包括相关文献的综述。一般采用两种形式,它们是由学位论文格式的文献综述演变而来的。第一种是两章节模式,首先是一个介绍主题的短章节(通常是第一段),由研究主题、简要回顾、要解决的问题以及关于结果的假设共同组成。第二段提供了评论性的、综合的文献回顾,要长一些。这一模式的难点在于第一段里如何帮助读者理解研究问题的重要性,因此文字必须能提供充分的信息但又不显累赘。

另一种更精简的形式是保留上面两段材料主要的内容,并把它们合并成一段。一般以一个对主题的简短介绍开始,然后着力呈现以如何产生研究问题和研究假设为主要内容的文献回顾。这一模式的一个不太常见的版本近似于一篇期刊文章的前言部分(4~8 页),有时还包括一份更长、更综合的附录。

不管你选择使用哪种形式,开题报告必须包含与研究主题相关的文献回顾。换句话说,写下一份将会在学位论文最终稿(也包括在你进行研究之后发表的最新研究)里引用的所有文献的综述。尽管委员会可能会接受"缩水版"的开题报告,但是我们强烈建议你提交一份完整的开题报告。我们之所以提出这样的建议,是因为你应该完成与写作开题报告相关的所有工作,否则就不会知道是否有人已经抢在你之前研究过这个问题了。一个完整的文献回顾也可以使你了解该研究领域的进程和设计细节,避免犯其他人的错误。写作一份精密的文献回顾将迫使你比仅处于阅读文献的阶段更加注重对细节的深入思考,并在开题报告中总结出简要的研究概述。

另外,如果已经完成了文献综述和方法部分的写作,那么你的学位论文就已经写出了一半。如果把文献回顾留待以后再写,那你还要在整个研究项目结束的时候做这项工作。而收集数据的时候你不大可能重新阅读这些你已经忘了的文献,所以工作将会一拖再拖。

在文献回顾中,我们同其他人(如 Yates,1982)一样推荐使用"漏斗法"原则:从所选的研究领域的一般文献开始查阅,然后一点一点地缩小到研究的具体领域,最后精确到你要研究的问题。介绍完主题和回顾完相关的文献之后,你要引导读者领会自己的研究领域的基本原理。假如你只写一份单一章节的文献回顾,就像之前提到的那样,你需要在文章结束的时候交代出研究的问题和假设。

文献回顾的篇幅应该有多长?以我们的经验来说,其中的差异是很大的,这些都取决于学院的要求以及你总结的相关文章的数量(例如,文献回顾的深度和你总结的背景资料的多寡)。我们认为博士学位论文的文献回顾部分的长度范围应该有 20~50 页,硕士的论文要少一点。

方法

如果以烹饪书来类比的话,在你把文章的前言仔细地完成的时候,读者就知道了为什么要把"巧克力慕斯蛋糕作为你晚餐聚会的点心了"。你选择搭配"主菜、酒和其他食品"的逻辑方式和组织安排的效果是如此恰当,此时没有人还会想着去挑选"其他菜式"。那么现在就开始来"烹任"吧!研究方法部分就是开题报告中关于"将如何完成研究"的内容。这部分内容应该提供详细的表述,这样任何读到的人都可以在需要的时候重复你的研究。也就是说,每个人在阅读了菜谱之后制作出来的蛋糕都应该是相同的。

研究方法一般包含着若干小部分。第一个小部分经常用来说明**被试**的特征。随后的部分包括实验设计、自变量、测量(因变量和其他)、仪器的设置和程序等。具体包含哪些东西,以什么顺序呈现则要由你的研究来决定。第 7 章对开题报告中的研究方法部分应该包括哪些内容提供了更详细的说明。

开题报告的长度应该是多少呢?这个还是要根据你想进行多详细的报告,研究的过程和实验的设计有多复杂,方法和附录部分加入了多少内容而定。我们认为大多数博士学位论文的研究方法部分一般为 15~25 页,硕士的自然要少一点。

结果

结果部分用来展示你得到的最终成果。以包含初步分析的一些小的分析内容为开端——例如,检查测量的信度,进行效度分析,检查分数的分布,等等。然后,描述用来检验假设和研究问题的统计过程。你可以用下面的方式来表示焦虑和任务难度在被试错误答案上的作用,如"进行 2×2 协方差分析,把学分积点作为协变量"。这表明你将会提供每一组被试在因变量上的平均数和标准差数据组成的表格。如果使用方差分析(ANOVAs)或者协方差分析(ANCOVAs),还需要提交一份总体(分数)的表格表示主效应、交互作用、误差和自由度,以确保使用的统计分析方法准确合适。读者也需要准确地知道研究中有多少个计分变量,以及测量的对象都是什么。本书第 9 章为统计方法和研究问题的匹配提供了指导。

在大部分的学院里,开题报告中的数据分析计划要么放在研究方法部分(作为开题的统计分析部分,见后面的专栏5.1)的最后,要么作为一个简单的“结果”部分进行呈现。其实开题报告中的数据分析可以比以上两种情况更为复杂一点,给答辩委员会清晰地呈现一个研究结果的合理猜想,比如应该为每一个相关测量组的平均数和标准差提供合适的表格。还可以将发现的结果和效应使用图形或图表来呈现,例如做方差分析时可以用图形来表示期待获得的交互作用。对结果的预期可以帮助你弄清楚究竟想要做什么,还可以方便日后检查文稿中的错误。如果设计了单一事件的重复测量或多组被试的重复测量,你需要提供适合这些数据的图表。如果研究的是描述性的问题并且包含对一定数据的相关测量,就出示可能得到的相关矩阵。

用正确的形式呈现结果——展现你预期的样子以及如果研究失败了的话它们会是什么样。最后,展示最有可能遇到的情况——即,正、负混淆的结果。

你可能想知道我们为什么要在学位论文开题报告的建议中提到对结果部分的要求。大多数人并不会提出这么细致的要求,那么为什么要做一些你不需要做的事呢?而且,你眼下的**目的**是做研究。你还没开始做呢,所以在这个部分中哪里有结果啊?

如果你把这部分内容当成“假设结果”或者是“预期得到的结果”,那就该感谢我们的建议。尽管不是所有人都同意在开题报告中加入这些内容,我们和斯腾伯格(Sternberg,1981)的观点相一致,那就是这部分内容会迫使你去思考如何分析和呈现数据。当你需要预测结果时,你将仔细地考虑需要使用的统计方法和过程。这些假设或预期的结果会帮助你发现数据是否符合统计假设,是否可以用来分析并回答已提出的问题,以及在研究的过程中还需要哪些辅助资源。从不同的角度看待可能出现的结果还能帮助你知道研究设计的弱点在哪里并且在开始研究前就做出更正。

预测研究的结果还可以帮助你提高自己的统计水平。例如研究需要用到路径分析、潜变量增长模型或者是因素分析这些统计方法,你就需要多学习这方面的知识。你也可以通过使用合适的软件来完成这些数据分析。另外,在研究数据可用时,你可能还需要就数据的分析问题咨询一下统计方面的专家,这可以在日后节约大量的时间。而且在收集好数据之后,你也可以立刻开始分析而不用停下来考虑使用何种统计方法。记住,在复杂乏味的研究上保持持续工作的动力尤为重要。另外,预期结果只是为了给繁忙的委员会成员提供你将要做什么研究的参考材料。通过这些额外的研究细节,他们可以知道研究存在的潜在问题并且在状况出现之前就帮助你解决。

文献目录

文献目录是开题报告中引用文献的目录清单。除非你的学校要求使用其他格式,否则请一律使用《Publication Manual of the American Psychological Association》[①] 要求的格式来准备这份清单。确定只列出了引用的文章,因为一些新手的清单包括了所有他们读过的文献,尽管这可能使答辩委员会对你的工作量刮目相看,但是这

① 本书中文版《APA 格式——国际社会科学学术写作手册》已由重庆大学出版社引进出版。——编者注

种清单是完全不必要的,一定要从清单中删除读过但没有引用的文献。文献目录可以在研究完成之后,或者在准备提交最后一版论文时再添加。

附录

附录是为将来想要重复你研究的人提供的包含所有额外细节的"仓库"。也是使答辩委员会成员确定你有能力完成研究的重要信息。被试结构、知情同意书、叙述脚本、研究中使用到的不涉及版权问题的收集信息的工具、结构图、仪器连接图,以及操作手册的复件都可以在开题报告的附录中找到。一般来说,有版权保护的测量工具不能包含在附录里,除非你从版权所有者那里取得了授权。

按顺序把以上的内容加入附录,这样读者就可以轻松地找到它们。因此,如果你第一次接触附录制作,那么请记住附录 A 应该包含一份刊登在当地报纸上的被试招募广告;附录 B 包括其他的材料等。

内容表

一旦完成开题报告的主体部分,你就可以开始制作内容表了。实际上,最好是以提纲的形式尽早开始准备,这样可以让你更好地组织文献回顾和研究方法部分。也可以避免使提交的材料过于概念化。提前准备内容表也为你应该如何完成学位论文提供了一个任务分析清单。一个仔细准备的三水平提纲或是内容表将清晰地呈现你需要在研究中完成的那些任务(见专栏 5.1)。在完成不同的任务时,可以检查一下自己已经完成了哪些工作,以从中获得巨大的满足感和动力。

当然,在整个开题报告没有完成之前,内容表是不会有页码的。实际上,这份内容表应该被视为修改写作进程的试验性提纲。当你准备加入页码(例如,就在你马上要把你的开题报告提交给委员会之前)的时候,确保已经将附录全部命好了名。例如,"附录 A:被试指导语……86 页",要准确地表明去哪里可以找到这些提供给被试的信息。不过也有一些学位论文仅仅标明附录 A 在 86 页,而不是指明附录里都有哪些内容,我们认为前者更好些。另一个比较好的尝试是在附录中把在正文中涉及的部分标出页码。因此,"告知……在每一个部分实验前可以获得 10 分钟的放松训练(见附录 A,86 页)"比仅仅出现"……(见附录 A)"的形式对读者会更有帮助。

专栏 5.1 续显示了一个学位论文开题报告的内容表。研究导师推荐给你作为模板的附表有助于你组织自己的开题报告。

调查一般的开题过程

学生们在不同导师的指导下准备开题报告的过程会有所不同,但其中也存在一些共性。当你:①和导师确定了一个研究问题(不一定按这个顺序);②汇总和阅读了相关的文献;③对将要做什么研究有一个好点子的时候,这个过程就开始了。然后你就可以开始写作了。一些学生会把他们已经写的材料交给导师(如,首先是文献回顾,然后是方法部分)阅读,在全部完成之后再提交给其他人看。你的导师会以他/她自己的方式来给你提供建议。

专栏 5.1

学位论文开题报告内容表样本

儿童攻击性评估:对于自我和同龄人报告方法的拓展

内容表

专栏5.1 (续)

注:海瑟·欧诗(Heather O'Shea,2003)的内容表。已获作者同意引用。

一般来说,除非有人了解你的研究领域或是对你的研究特别感兴趣,否则不会仔细阅读你的早期草稿。导师应该对你已经写出来的东西给出反馈,然后你要反复修改,直到他/她认为可以提交给答辩委员会成员为止。通常,将报告提交给答辩委员会就是你准备好了参加开题报告会(见本书第13章)的信号,那时你的读者将提出需要进一步思考的问题并且建议你对文章内容做出怎样的修改。开题报告会之后,你需要按照委员会的要求和其他的附加条件修改、提交开题报告。

在此程序中的有些地方你还要准备和提交附加的文件来获得机构审查委员会(IRB)对于使用人类(或动物)被试进行实验的许可。作为获得联邦基金的条件,美国健康和人类服务局要求代理部门保证进行实验的个体可以获得报酬。不管是在大学、医院、诊所,还是社区,只允许进行那些由 IRB 或者人权组织进行监督的研究。

当地的 IRB 由联邦政府管理并且遵循当地法规和传统办事。在你开始收集数据之前,必须获得当地 IRB 的许可才能进行研究。你需要准备的文件应该提供了对研究过程的概述、对伦理问题的处理(如,保密、程序风险)和可能存在的任何与研究有关的具体的伦理学争议(如使用欺骗)问题。我们在第7章对研究的伦理问题有更多的阐述。

不同的学院和导师对以上程序的时间安排和要求也不同。例如,本书的一个作者就要求学生只有在获得 IRB 许可的时候才能提交开题报告。另一些导师则相反,他们要求学生只有在获得答辩委员会正式同意开题之后,学生们才可向 IRB 申请批准研究!向其他的学生咨询一下,查一查你所在学校对开题的要求,或者问问你的导师有哪些关于这些程序和时间安排的要求。

思考与练习

写作开题报告

☐ 理解学位论文开题报告的重要性

—明确学院或学校的要求

—开题报告是否需要装订

☐ 了解学位论文开题报告的内容

—内容表

—前言（文献回顾）

—方法

—结果

—文献目录

—附录

☐ 了解一般的开题报告程序

—时间表

—导师对草稿内容的偏好

—什么时候提交给委员会成员

—什么时候以及如何提交获得 IRB 许可的协议书

—学院的要求

6

文献综述

当确定了研究的主题之后，你需要获得更多与该主题相关的信息。这就需要查阅和总结与主题相关的专业文献，阅读并且思考前人都有过哪些研究。这个过程不仅可以帮助你认识该领域的核心概念和方法，也能够使你了解到是否有其他聪明的研究者已经捷足先登。此外，一个间接的好处是在接下来的数个小时里你尽可以放下手中的笔，安心地坐在扶手椅上享受阅读的乐趣。

查找相关文献

与你的研究相关的文献资料通常来自学术期刊、书籍以及互联网。除此之外,会议论文集、尚未发表的论文和博士学位论文也可以是好的文献资料来源。

上述每一种来源之中都包含了大量的信息,那么怎样才能有效地从庞杂的信息之中挑选出与特定主题关系最为密切的文献呢?一些传统的或是高科技的查找方法为检索相关文献提供了便利,虽然每种方法都有其自身的局限性,但如果可以将它们结合起来使用,就可以确保你完成一次全面合理的文献检索。在开始文献检索之前,这里还有两条重要的建议:①依靠原始资源;②避免使用大众媒体发布的信息。

使用原始资源

请找到原始文章(即原始资源)并亲自阅读它们,而不要仅仅依赖于其他人对某一研究所做的描述或总结(即次级资源)。这是因为评论者经常会错误地将原著没有提到过的事情煞有介事地加以表述,甚至你还会哭笑不得地发现自己的研究被援引作为相反立场论点的佐证。除了以上原因外,你可能根本就不同意其他评论者的结论,那么现在正是展现聪明才智和表明自己观点的时候。当然,这并不是说其他人对原始文章的总结或评论都是毫无价值的,你可以尝试着在亲自阅读过原文之后再去看别人的总结,进而思考自己是否赞同这些次级资源的观点。

你可能会问这样的问题:“有的时候我确实无法获得原文啊!一些文章会发表在偏僻难查的期刊上,或是作为大会论文就从来没有被公开发表过。”在这种情况下也不要直接引用你没有读过的文献,而是应该依照美国心理学协会(APA)的规则标明你所引用的文献来自于次级资源(例如,“Foster,2000,cited in Cone,2005”)。

避免使用大众媒体发布的信息

尽管诸如时代周刊(*Time*)、连线杂志(*Wired*)、纽约时报(*the New York Times*),以及滚石杂志(*Rolling Stone*)这些我们经常阅读的刊物可能是知晓当今时事的捷径,但是它们仍无法取代经过同行评议的期刊中学术文章的地位。对于网站来说也同样如此,一些网站信息的来源和准确性都是没有保障的。尽管大众媒体上的信息也可能是准确的,但因为这些信息没有经过严格的审查,所以往往很难对它们的可靠性做出估计。

需要强调的是,由于网络资源具有多样性的特点,所以需要慎重地对待网站上的信息。一些网站是由对某些问题感兴趣的业余爱好者建立的,他们主要是想借网站抒发自己的观点;而另外一些网站则是由严谨守信的科研机构创建的,它们提供的信息多是由一组科学家基于新近的文献汇编而来。由此可见,对互联网信息的评估是十分关键的:网站是由科研机构或专家创建的吗?你能够确认信息的来源及其作者吗?获取信息时,请务必选择能够以肯定形式回答上述问题的网站,至少信息的来源应该是可靠的,并且信息本身也应该是能够令人信服的。

确定主要研究者和重要的期刊

收集某一研究领域的相关文献要从哪里开始动手呢？你可以先确定在这一领域中都有哪些著名的研究者，他们通常都会将自己的研究成果发表在哪些刊物上。我们将在下面介绍3种获得这类信息的方法。第一种方法最简单：你的导师或同学可能对你所研究的领域有更深入的了解，所以不妨去问问他们对收集信息有什么建议。

另一种好的信息来源是相关主题的最新书籍或篇章的参考书目，你可以通过浏览它们找到相关的文献以及这一领域著名研究者的信息。同时，也可以通过阅读这些书籍和篇章进一步了解这一领域中关键的问题和概念。

第三种高效的方法是，用 Google 找出重要研究者的信息。打开浏览器进入 http://www.google.com/然后输入研究者的姓名，就可以找到这一研究者在相关领域中最新的信息和进展了，很有可能你还会从中发现许多尚未正式发表的研究成果。除此之外，你还可能借此方法找到研究者在其所属大学网站中的个人网页，发现研究者最近一段时间在相关领域所做的演讲或研讨会的信息。当然了，除了 Google 之外，你也可以利用诸如 Yahoo(http://www.yahoo.com/)、Alta Vista(http://www.altavista.com)等其他搜索引擎做上述检索。最后，http://www.guru.net/也是一个值得利用的网站，这个网站几乎包含了所有你能想到的学术话题的相关信息，通过它可以获得许多有用的资源。

使用参考文献索引

许多参考文献索引能够对文献检索工作有帮助，在这里我们将介绍它们之中比较有代表性的几种。《心理学文摘》(*Psychological Abstracts*)：该索引中包含了大量心理学学术期刊的论文摘要，这些论文摘要依据其主题和作者进行分类。心理学文摘是心理学文献数据库(PsycINFO)的印刷版，一般可以通过大型图书馆的参考咨询部获得。《国际学位论文文摘》(*Dissertation Abstracts International*)：该索引与《心理学文摘》(*Psychological Abstracts*)类似，专门用于检索学位论文文摘资源。科教资源信息中心(Educational Resources Information Center)：在这里能够找到大量与教育相关的信息。《期刊题录快讯》(*Current Contents*)：该索引每周更新一次，它会列出新近期刊目录的一览表，这能帮助读者确定向谁索要再版的信息。最后，值得一提的是美国心理学会出版的一系列不同主题的 PsycSCANS，它们根据各自的主题提供相关领域学术期刊的文献摘要，例如 Developmental PsycSCANS 中所包含的杂志一般都是与儿童发展心理学相关的。尽管期刊题录快讯和 PsycSCANS 很难在图书馆中查到，但是你所在的心理学系可能会订阅它们，你可以借来复印。Cooper(1998)对这个话题做了更详尽的介绍，并且还提供了一个如何运用各种参考文献索引的概述，如果你对此感兴趣，可以找来看一看。

我们在这里要特别提及一个未被充分利用的文献索引——《社会科学引文索引》(*The Social Sciences Citation Index*, *SSCI*)：该索引包含社会科学中50多个不同学科、1 500多种学术期刊的文章信息。它的主要功能是帮助你了解在一篇文章发表后，都有哪些人在之后的研究中引用了这篇文章。通常你可以获得社会科学引文索引电子版的 CD，也可以在网络上的 Social SciSearch 中找到它。如果想要使用这个索引，需要知道原始文章的题目是什么。当你想要弄清一篇文章发表之后研究又有

什么新的进展时,社会科学引文索引是一个十分有用的工具:你可以通过它知道一篇文章对其后续研究的影响,并且还可能发现一些利用计算机数据库检索可能漏掉的参考文献。此外,当你决定使用某种心理学的测量工具时,社会科学引文索引也可以帮助你找出包含该工具信息的研究,多数测量工具的使用者都会在自己的文章中引用有关此工具的最早的或是最富开创性的论文。许多同学发现,如果没有人来指导是很难掌握社会科学引文索引的使用方法的,所以在使用这个索引之前要仔细阅读它的说明资料,当你第一次使用它时最好向图书馆参考咨询部的馆员请教一下该怎么使用它。Cooper(1998)指出,使用社会科学引文索引可以弥补一些由于期刊参考文献中存在错误而带来的损失。为了确保自己已经找到大部分相关的参考文献,他的建议是,务必要找到第一个从事这项研究的作者的文章,尽管这样做会使你面对一个很长的参考文献列表,但它却是全面的,你只要再剔除掉那些无关的文献就可以了。

除此之外,一些心理学中特定领域的参考文献索引,甚至是心理学以外其他学科的索引也可能会对你有帮助,你可以去那些大型的大学图书馆,那里的馆员会帮助你找到需要的文献索引。

运用计算机检索文献

计算机数据库能够使你更加快捷地进行文献检索。如果想熟练地运用各种计算机数据库,你需要掌握如何进行检索的知识。首先,找出关键词(keywords)或是研究术语——你觉得能够描述文章主旨的单词或短语。许多期刊都会在论文第一页的摘要下面列出关键词,这使得找出关键词这项工作变得很容易。接下来要做的是,将找到的这些关键词输入到搜索引擎中,并且告诉它到哪里去查找这些单词,计算机会按照你的特定要求到摘要、题目或是文章中进行搜索。通过作者姓名、期刊名称的检索方法与通过关键词检索的方法是相同的,在这里不加赘述。需要说明的是,你可以让搜索引擎查找一篇含有多个关键词的摘要(列出几个关键词,用逻辑关系词"and(和)"连接它们,之后计算机会找出同时包含所有关键词的摘要)。当然,你也可以命令计算机找出含有任意一个(不用所有都具备)关键词的文章,这个时候几个关键词之间的逻辑关系词要用"or(或)"。获得检索结果后,既可以在屏幕上浏览检索结果,也可以把结果打印出来,把它们下载到存储设备中,或是将它们通过电子邮件发给自己。一些搜索引擎可以将你感兴趣的信息自动发送到你的电子邮箱之中,十分便利。最近几年被人们广泛运用的博客和播客也可以是一些有价值信息的来源,但现阶段它们更多的还是为那些研究兴趣相投的研究者提供一个分享各自观点的平台,而非具体准确信息的来源。如果你还想对博客和播客有更多的了解,请查阅维基百科(http://en.wikipedia.org/)。

一个利用计算机数据库检索的误区是,以为用这种检索方式就是对文献进行了一次全面的搜索。其实未必如此,因为你所做的检索:①只是对你选择的特定关键词的搜索;②只有当作者在题目、摘要或文章正文等处用到了与你选定关键词一样的词汇时,文献才可能被查找出来;③只能查找出数据库包含的期刊中的文献;④只能查找到在数据库包含的年限中发表的文献。

由此可见,找到恰当的关键词是文献检索成功的关键,这并不是一件容易的事情,因为并非所有的研究者都用一模一样的术语描述自己的研究。比如说,我们最

近的一项研究用到了 trichotillomania(拔毛发癖)这个词,而其他的研究可能会使用 hair twirling 这个词代替 trichotillomania 去描述同样的概念。因此,最好先用一般性的词汇作为检索的关键词(如 hair twirling),而不是让你的检索一开始就受制于冷僻的专业词汇(如,trichotillomania)。

如果期刊没有单独提供关键词,那么有什么其他好方法能找到合适的关键词呢? 可以找来一些与你想要研究的主题相关的文章,通过它们的摘要确定关键词,一般来说反复出现的词汇可以作为关键词的备选对象。许多计算机数据库都具备一个十分有用的功能,即术语的同类词汇编。这个功能可以帮助你挑选出关键词,同时也能帮你复查一下你的项目中是否存在漏掉的主题。此外,你还可利用同类词汇编功能确定自己文章的关键词。

如果你对自己研究领域的计算机数据库不是很熟悉,可以先到 http:// researchresources. net/directory/socialsscience 看一看,这个网站提供了上百种与行为科学及教育科学相关的网站的网址。此外,也可以浏览一下 http:// www. socsciresearch. com/。

在心理学研究中,心理学文献数据库(PsycINFO)可能是最常用的计算机数据库了。这个数据库中涵盖了自 19 世纪 80 年代以来的心理学文献摘要,其中包含有期刊(1 900 余种)、书籍、篇章、学位论文摘要以及其他数据库。如果想要查找全文可以去心理学全文期刊数据库(PsycARTICLES™),这是一个与 PsycINFO 相链接的数据库,它包含了美国心理学会和各类其他团体从 1988 年到现在出版的 42 种杂志、31 000 多篇文章的全文资源。交付一定数额的服务费,就可以获得 PsycARTICLES™的资源,当你急需获得一篇重要的文献时,这个数据库可以为你提供极大的便利。前文中提到的 PsycSCANS 原本是一种纸质的文本数据库,而现在的 PsycSCANs Online 已经成为了 PsycINFO 的一部分,它主要包括参考文献引文和各类特定主题(如应用心理学、发展心理学)的摘要。

期刊题录快讯数据库(Current Contents Connect®)是前文中提到的期刊题录快讯(Current Contents®)的网络在线版。如果对与健康主题相关的文献感兴趣,你可以去检索一下医学文献数据库(PubMed http://www. ncbi. nlm. nih. gov/entrez/),这个数据库是美国国家医学图书馆(National Library of Medicine)提供的一项服务,它包含了 20 世纪 50 年代至今超过 1 400 万篇生物医学论文的引文。此外,还可以从医学文献数据库中获得很多全文或其他相关的资源。另一个能够获得全文资源的地方是 Ingenta(http://www. ingenta. com),在我们撰写这一部分的这一时刻,你可以从中找到来自 28 618 种出版物的 16 719 850 篇文章,尽管这是一个文献检索的好地方,但是这里的全文资源价格很高,你如果能够通过图书馆去购买这里的文章能够相对便宜一些。

最后,谷歌学术搜索(http://scholar. google. com/)也是一个免费的在线搜索引擎,它可以帮助你找到相关的文献资料。

索要文献

当确定了在你的研究领域中有重要贡献的研究者,你可以通过写信、电话或电子邮件向他们索要文献,随着电子文档与电子邮件的出现和发展,索要文献已经变成了一件十分便捷的事情。多数作者是很愿意向你提供期刊论文的,当然一些作者会收取一定数额的影印费或邮资也是正常的事情。此外,你还可以索要那些尚未刊

登在期刊上的篇章,这些篇章可能已经是待印版(已递交或被接受但尚未正式刊发的论文)或是即将发布的大会论文,一旦得到了这些文献,你便可以更早地了解到研究的新动向。但是,我们也要提醒你,并不是所有作者都会满足你的要求向你提供某些篇章的复印版、待印版或学术演讲稿。你可以利用我们之前提过的搜索引擎获得作者的联系方式,或者也可以从一些相关组织的成员名录中得到这类信息,如美国心理学会(American Psychological Society)、儿童发展研究协会(Society for Research in Child Development)或是行为分析协会(Association for Behavior Analysis)。你往往可以从一些学校或机构的图书馆以及学校心理学系主任那里获得上述成员名录。在这里需要注意的是,如果想很快得到回复,打电话往往比写信更加高效,因为比起一沓不起眼的信件,想要忽视或拒绝正在和你通话的人的请求往往要更困难一些。当然,如果能用电子邮件的话,效果可能也会很好。

Yates(1982)指出,文献检索者还可以查阅一下相关的大会论文,许多研究成果会在学术会议之后发表于大会论文集,你也能从中了解到研究的发展动向。

查看其他学科的相关文献

《心理学文摘》、心理学文献数据库和其他一些资源自身涵盖的刊物数量是有限的,所以能够检索到的资源也是有限的。你会发现自己的研究可能与其他学科有重叠之处,比如在医药学、通信学、教育学和社会学中都包含一些与心理学相关的期刊或文章。如果你从事跨学科的研究或者你的研究与其他学科有交叉,可以去检索一下其他学科的数据库,或者去找一下那些学科的参考文献索引。

浏览核心期刊目录

你已经开始熟悉自己的研究领域,下一步要做的是去查一查该领域中多数文章都发表在哪些期刊上,然后浏览一下这些期刊近十年来的目录一览表(通常列在每卷的结尾处),此举能够帮助你挑出那些由于是刚发表而还没被收入参考文献索引或数据库的文献,而且这也是一个帮助你重新检验你利用其他方式查找的文献是否全面的过程。除了刊物本身外,在期刊题录快讯或是美国心理学协会的网站(http://www.apa.org/journals)中也能找到一些期刊目录的一览表。

利用参考文献列表

当你查阅相关的文献时,关注一下它们提供的参考文献列表并记录其中对你有用的相关文献,这也是一种文献检索的好办法。Cooper(1998)指出了同时利用核心期刊和浏览参考文献法的缺陷,他认为大部分学者在某一特定领域都有其非正式的人际关系网络和期刊网络,他们往往倾向于在某一特定网络中发表自己的文章或引用他人的文章,这就导致了这样的结果:对于该网络内部的文章引用过度,而对该网络之外的文章却引用不足。所以我们鼓励大家使用多种检索方法来全面地查找相关文献。

如果你按照我们的建议完成了上述步骤,并且发现自己在各种不同的参考文献列表和索引中找出了相同的文章时,那么你的文献检索工作就已经基本完成了。尽管这并不意味着你收集到了某一主题涉及的全部信息,但至少你已经找到了这个领域之中大部分最主要的论文。

记录引用条目

浏览相关论文的参考文献列表时，可以把摘录的整个引用条目（以 APA 格式）写在索引卡片上（每张卡片记录一条引用），或者是将它们存为电子文档。我们更推荐后一种方法，尤其是你带着一台笔记本电脑或是一台掌上电脑（Personal Digital Assistant，PDA）去图书馆时，后一种方法显然更加便捷。养成一鼓作气把引用条目都敲进计算机的习惯，并认真去学习最后呈现在参考文献列表中的引用条目应遵循什么样的规则。对不同类型的引用条目有不同的 APA 格式要求。我实在不愿意回忆自己第一次写参考文献列表时的痛苦经历，因为没有及时地记录，导致我不得不耗费大量的时间再重新去查找那些引用条目。一些同行强烈推荐 EndNote，这是一种可以帮助你管理引用记录和设定各种引用模版（包括 APA 格式）的参考文献数据库。但是我要在这里提醒你：即便最优秀的软件也不是百分之百可靠的，你要为引用条目的精准性负责，因此，还是奉劝你亲自去学习一下 APA 格式为好。

当我们记录将要查阅的论文时，可以将参考文献的条目打印出来。如果在一张 81/2″×11″①的纸上打印 3 个条目，可以把这张纸裁成 3 条。它们的作用跟索引卡片类似：卡片上面的条目能够帮助你在图书馆中找到相关的文章或书籍，之后还可以将“卡片”插入期刊或书籍中以标记需要影印和做笔记的页码。此外，也可以将这些卡片按照一定的规则编排起来，从中找到想要收集的特定信息，再将它们输入到电子表格中，在需要的时候，它们就可以按照字母表排序，成为你的参考文献列表了。需要注意的是，并非所有出版物上的引用条目都是精确无误的，所以无论你将要参照怎样的格式，在录入参考文献条目的时候，一定要确保它们是依据原始研究的文献信息精准记录的。

批判性阅读

完成收集文献的工作后，就该对它们进行批判性的阅读了。记住，你要做的工作是将这些信息综合起来形成一个条理清晰的综述评论，这个综述评论应指出研究的主题、优点和不足。为了实现这个目标，应该具备一种既不忽略每棵树木又能着眼于整片森林的能力。那么到底该如何把庞杂的论文组织成一篇文献综述呢？一个不错的办法是，浏览一下你收集到的论文或篇章（Galvan，1999），尝试着找出这些作者所运用的方法论和概念框架的主题，以及它们之间的共性。在这里需要提醒你的是：不必引用每一篇阅读到的文章！有哪些文章与你所研究的领域关系最为密切，哪些只是有一点点的关联？有哪些主题是真正需要运用的？又有哪些是可以跳过的？带着这些问题来阅读文章，并记录下每个研究在方法上的优点和不足，这些工作将使综述的写作变得更加便利。专栏 6.1 提供了一系列可以用来评价实证研究的问题，也可以参看 Galvan（2004）的作品来获得其他方法。

对于实证研究的论文，要仔细阅读它们有关理论依据的阐述（通常写在方法部分前面一个段落中）、研究方法和研究结果这 3 部分，至于研究的结论部分我们不赞

① 216 mm×279 mm。——译者注

成直接阅读,因为你应当具备自己得出结论的能力,将你的观点和作者得出的结论加以比较,在这个过程中,你可能会发现一些作者没有想到的解释,或是发现作者其实过高地估计了自己研究结果的价值。

专栏 6.1

实证研究评价指南

引言

1. 引言部分是否提供了研究的理论依据?
2. 是否清楚地交代了研究问题和研究假设?(请注意研究问题通常隐藏在对研究目的的描述中。)

方法

1. 是否对研究方法做了详尽的描述?
2. 被试
   a. 是否对招募和选择被试的方法进行了说明?
   b. 被试的选择是随机的吗? 抽样时是否可能存在某种偏向?
   c. 样本是否能够代表总体?
   d. 是否对样本的特征进行了充分的描述?
   e. 如果是将两个或多个群体进行比较,那些潜在的且有可能令人混淆的变量是否具有可比性(如人口学变量)? 如果没有,这样处理是否合适?
   f. 是否向被试介绍了实验或研究的具体情况,并征得其同意?
   g. 样本容量足够满足测量数量的要求和发现效应的要求吗?
3. 实验设计
   a. 如果需要,是否采用了控制组?
   b. 控制得是否合理?
   c. 控制的目的是什么?
   d. 如果是一项实验研究,被试是否被随机分派到各组之中?
4. 测量
   a. 对于所有的测量(例如,用于归类的测量或是因变量等),研究者是否给出了它们的信度和效度信息?
   b. 信效度信息是否说明研究所采用的方法是合理的? 当某一种测量方式第一次出现在某种研究时,这一点尤为重要。
   c. 测量方法是否与研究问题和研究假设中提到的相符?
   d. 如果用到了不同的任务或测量方法,它们的顺序是否得到了平衡? 作者是否对潜在的顺序效应进行了分析?
   e. 是否运用了多种测量方法? 尤其是采用不同的方法(自我报告法、他人评定法、自我监督法或直接观察法)对那些来自同一领域样本进行研究。
   f. 如果研究中涉及观察、判断或者评分者评定,那么是否对观察者或者评分者之间的一致性(信度)进行评估? 数据是否来自于具有代表性的样本? 两个评分者是否独立地进行评定? 他们的评分信度是否令人满意?

专栏6.1 （续）

5. 偏见和人为因素

a. 研究的执行过程和测量的评定是否在双盲的情况下进行（例如，研究助手不需要知道研究假设，以避免产生偏见或人为因素的干扰）？

b. 如果是一项准实验研究，那么在研究设计中作者是否包括了恰当的步骤，以控制与研究结果可能会有争议的地方？

c. 在所有的组别中，针对全部被试的操作过程是否具有一致性？是否说明了如果运用不同的操作可能会引起一些混淆？

6. 自变量

a. 如果是一项实验研究，那么是否有这样一项检查，即实际操纵的自变量正如研究设计描述的那样在执行？

b. 如果是一项干预研究，是否有足够的心理治疗干预专家执行干预（如为了增加干预的效果）？

c. 如果是多个处理水平或条件进行比较，作者是否阐明了这些条件在方法上的不同，在其他方面是否相同（如，时间的长度、干预者的资格）？如果存在差异，那么这种差异是否会严重地影响研究结论？

d. 实验程序和自变量在哪些方面限制了研究的外部效度？

结果

1. 数据是否满足统计计算的前提要求（如拟作重复测量的方差分析的几组数据方差是否同质）？

2. 是否恰当地使用了显著性检验，并且对其进行了充分地报告（即是否为接下来的分析工作提供了足够的细节）？

3. 在相关研究中，作者是否指出在不同的方法之间，存在着大量的共享变量，那么作者是否解释了为什么有的相关程度低，但是却具有显著性？每一种方法都有其局限性，而这些相关是否受到测量方法的限制？是否因为作者提供了平均数和标准差，所以你能够判断结果？

4. 如果进行多次的统计检验，作者是否调整了 α 的水平或是使用多元统计分析方法来降低这种情况下发生 I 类错误的可能？

5. 作者是否报告了平均数和标准差（如果需要），使得读者能够检验有显著差异的结果是否有意义？

6. 对于多元统计分析，被试数目是否符合最低要求（每个因变量至少 7 名被试）？

讨论

1. 作者是否如讨论差异显著的结果那样对边缘显著或不显著的结果进行了讨论？

2. 作者是否对数据做了过度的解释？（例如，将相关关系的结果表述为因果关系，或是如解释直接观察的结果那样解释行为的自我报告结果。）

3. 作者是否考虑到结果可能存在的另外一种解释方式？

4. 作者是否指出自己研究的局限性（包括方法上存在的问题）？作者是否指出研究中被试、程序和因变量等因素限制了其对研究结果的解释？

5. 作者是否“接受”了虚无假设？

注：转自“A Reader's, Writer's, and Reviewer's Guide to Assessing Research Reports in Clinical Psychology,” by B. A. Maher, 1978, *Journal of Consulting and Clinical Psychology*, 46, pp. 835-838.

将你在阅读过程中的想法和观点记录下来,当这些想法和观点能够描述大部分文献的特点时,你对研究主题就有了一个更为深刻的认识。另外,在阅读收集到的文献时,请将它们按类别分组。

保存论文的全文电子版或打印版,或者——如果资金有限的话——也可以只保留你的阅读笔记(或记录卡片)。我们更推荐前一种方式,因为尽管前者可能要花费更多的资金,但当你日后一旦要为了某一个细节而重新查阅已有资料时,保留下来的全文资源能够为你提供更为详尽的提示,而这些提示极有可能是阅读笔记中没有记录的。此外,如果你准备在这个领域有更深入的研究,这些文献的电子版或打印版还可以为你提供一个可以循环利用的数据库。如果你选择的是保留阅读笔记的方法,请一定要确保笔记中包含了研究方法的每一个细节:它们现在看起来可能意义不大,但一旦你尝试去解释互相矛盾的结果,或是去找出他人研究中潜在的问题时,这些细节将会发挥巨大的作用。一个好的方法是,列出一个需要在阅读过程中记录哪些问题的清单(例如,被试的数量和特征、自变量等),你还可以将清单发展成一个标准化的记录表格,我们提供了一位研究者在进行文献回顾阅读时所用到的记录表格,参见专栏6.2。这个标准化的方法可以帮助你避免在记录过程中遗漏未来可能需要的信息。(Cooper [1998] 提供了一个有关记录表格话题的扩展阅读)此外,一些软件程序也可以成为对你有帮助的资源,你可以利用文字编排程序(如 OpenOffice 或 Word)创建一个记录表格,并将它们存储在计算机或掌上电脑中以备后用。Filemaker™ 这样的数据库程序也能帮助你完成记录表格的制作和存储,而且还可以利用它从大量的资料中快速地找出特定类型的信息(例如:使用了遥感听觉提示的研究有哪些;采用6岁以下儿童为研究对象的研究有哪些)。

当在你的笔记中引用其他研究者的原文时,*务必把它们放在引号里,并写下它们的原始出处及页码*。因为事后很容易忘记哪些笔记是用自己的语言记录的、哪些是引用或演绎别人的话。即便你不是有意的,但是未经适当地引用声明就采用他人的观点或文字就是一种抄袭行为——我们将在本章最后再次强调这个问题。

最后,要避免过度阅读——一种常见的耽误时间的错误做法。在撰写自己的文献综述之前,将某一研究领域的全部文章都记住是没有必要的,为每一篇文章做冗长的笔记更是在耗费你的时间和精力。根据以往的经验,你需要以不同的方式阅读文献至少两遍:第一遍时,通过浏览它们来了解某一领域中的核心文章,并选择出哪些材料是与你的研究相关的,再想一想用什么方法来组织这些文章。在开始动手撰写文献综述时,你可能会意识到需要更加深入和细致地了解一些作者所做的研究及其发现,那么就找出之前选择出来的文献开始第二遍阅读吧。希望这个有侧重的阅读方法可以帮助你尽早地开始动手写作,永远不要将阅读作为你逃避写作的借口!

专栏 6.2

文章细节信息的记录表格

作者：________________

题目：________________

来源：________________

是否获得全文：□是 □否 □索要中

研究类型：□单个家庭 □多个家庭

问题行为：________________

直接涉及的儿童数量：________________

独立测量的类型和施测对象：________________

主试是否对行为进行了观察：□是 □否

主试实施观察的地点：□家庭 □临床 □实验室 □居民委员会 □学校

治疗（或训练）的次数：________________

设计：□ 无测量的案例研究

□ 附带测量的案例研究：□前测 □中测 □后测

□ 时序设计（000×000）

□ 相等时间取样设计：□A-B □ A-B-A □B-A-B □其他

□ 多重基线设计

□ 无测量的单个群体研究

□ 附带测量的单个群体研究：□前测 □中测 □后测

□ 单个群体的时序设计（000×000）

□ 单个群体的相等时间抽样设计：□A-B □A-B-A □A-B-A-B □A-B1-A-B2 □B-A-B □其他

□ 经典双群体设计（实验组和控制组）

□ 多群体设计：控制组、处理组和其他参照处理组

□ 两个或多个群体设计：其他参照处理组、经过处理的控制组

□ 因素设计

□ 其他：________________

后续研究：□是 □ 否

后续研究与先前研究的时间间隔：________________

简评：________________

____

写作准备

在开始写作之前，请牢记一点：你所撰写的论文不是用来自娱自乐的，而是供读

者阅读的。你可能对自己研究的内容很清楚,但论文答辩委员会的成员不能直接看到你大脑里的想法。他们不像你对文献知道得那样多,甚至有些人对你的主题根本不熟悉。因此,撰写文献综述的目的就是要让答辩委员会成员了解相关主题,同时也要使他们相信你是一个有能力的研究者:让他们看到你已经整合了所读过的材料并对信息的质量做出了评价。通过对文献综述的阅读,答辩委员会成员应该能够明确你所研究的问题、步骤和这个领域的标志性成果。他们还应从中了解到前人研究的不足、文献的空白以及发展该领域还需做哪些事情。等到他们读完文献综述,就应该能够清晰地了解到你的研究逻辑,以及你将试图在哪些方面改进前人的研究方法、设计和程序。如果综述组织得十分巧妙,你会引领读者得出这样的结论,即你正在做的研究必将是这个领域中最为杰出的。

了解文章长度和格式的规范

不同的学校对文献综述有各自不同的要求,因此在动笔写作之前一定去了解一下你的学校对文献综述的要求和规范。我们在第 5 章讲述的内容就可以作为对不同规范的应对方法:一个章节和两个章节的组织模式有着一定的区别,用其中不同的格式,你就会组织出不同形式的论文。

文献综述会因为要求的不同,在长度和格式上有很大的差异。你可以向论文答辩委员会主席询问一下文献综述需要的长度和全面程度:这将决定你如何定义自己的研究范围。再次强调,我们是在让你咨询答辩委员会主席而不是通过阅读其他文献获得文献综述的规格,这是因为每个学校对文献综述规格的要求都是不同的。比起撰写简洁明了的文献综述,写出一个四五十页甚至更长的文章其实是更加容易的,但事实上一篇 20 ~ 25 页的综述就足以看出一个学生在整合文献方面的能力了。和冗长而试图面面俱到但每样事情都没有交代清楚的论文相比,我个人更加偏爱简洁并有所侧重的综述,而且这种短小精悍的综述发表的可能性更大(许多期刊不会受理多于 25 ~ 30 页的文章),所以还是尽量将你的文章写得简洁明了为好。

写一个初步的提纲

为了对将要撰写的综述有一个整体的计划,你需要撰写一个写作提纲,在这个提纲中列出你打算在综述中用到的主要标题。在每个标题旁,注明你计划分配给每部分大约多少页。专栏 6.3 是一个学生的文献综述写作提纲,从它的页数分配上可以看出这位同学想要写的是一篇约 25 页的一章式文献综述。请注意:由于不同学校的标准不同,所以在写完写作提纲后,一定要请导师对文献的组织编排和页数分配给出评价。

专栏 6.3 这个提纲遵循了我们在第 5 章里讲到的漏斗法:从总背景出发,逐渐变得具体,最终落脚于对现有研究的评价,从而引出自己的研究的理论依据。

一个初步的写作提纲会让你了解到如果按计划来每个部分大概要写多长,以及是否需要限制或者扩充综述范围。即便你痛恨这种写作提纲,或是觉得它们毫无用处,那也请你将综述需要包括的内容以及页数分配列出来,因为它们能够帮助你知道现有的研究范围是否太过宽泛或是太狭窄。其实写一个提纲还是很有好处的,比如可以让同学或导师帮你看看这个提纲,从而获得关于文章的编排和逻辑是否清晰

的反馈；此外，写作提纲还可以减少因为前路漫漫所带来的焦虑和拖沓，看看提纲的页数分配吧！你会发现其实真正要做的事情不过是写几个5~10页的小论文（综述的每个小部分），再用一些过渡段落来将它们连接在一起而已，想想你的学习生涯中写过多少篇5~10页的小论文吧！这一次也一定能行！

如果你在正式的写作时发现需要修改提纲中的计划，不要感到惊讶，因为这是很正常的事情。就拿专栏6.3的作者来说吧，在正式写作的过程中，因为作者发现有关种族部分的综述将远远超出两页，所以她不得不将计划中的一部分砍去（学步期儿童部分）。把写作提纲当成是一种帮助你动笔写作和组织材料的工具，要懂得灵活地对它加以改动。

专栏6.3

写作提纲和页数分配

学位论文主题：9—12岁儿童的种族、性别和其处理与同伴冲突所使用的策略之间的关系

提纲

1. 引言（2~3页）
   - A. 以往有关同伴冲突的研究
   - B. 冲突在儿童发展进程中的重要意义
   - C. 对冲突的界定
2. 学步期儿童（6个月—3岁儿童；2页）
   - A. 冲突事件
   - B. 解决策略
3. 学龄前儿童（3—5岁儿童；5页）
   - A. 冲突的频率和长度
   - B. 冲突事件
   - C. 解决策略
   - D. 结果
   - E. 性别差异
4. 童年早期（5—9岁儿童；5页）
   - A. 冲突的频率和长度
   - B. 冲突事件
   - C. 解决策略
   - D. 结果
   - E. 性别差异
5. 童年晚期（9—12岁儿童；2页）
   - A. 目的
   - B. 策略
6. 种族（2页）
   - A. 比较研究

专栏6.3 （续）

B. 化解冲突的风格

7. 总结和评价(3 ~5 页)

8. 问题的提出和研究假设

注：Outline prepared by Avid Khorram, California School of Professional Psychology. San Diego, December 1991. 使用经过作者授权。

限定综述范围

在提纲上分配好页数之后,你可能会疑惑:我怎么才能把这么多的信息都塞进区区 25 页纸中呢？当有这样的想法时,你可能犯了两个错误。第一,你以为必须要将读到的所有材料都引用到综述之中,并且这个综述必须要包含所有和主题相关的信息,像是研究的历史、理论、新近的探索、他人的思考和奇闻轶事等。这是一种需要被纠正的看法,因为虽然你确实需要了解所有与主题相关的背景知识,并且能够回答关于它们的问题,但是无需在文献综述中将它们都交代出来,一个对特定研究领域的描述和评价就已经足够了。

顺便说一句,学生经常会问综述需要回顾多久的历史？可不可以将 10 年以前发表的文献都忽略掉？我们认为,在阅读文献的过程中,不管一篇文献是什么时候发表的,你都应该阅读文献的实证研究部分,藉此对这个领域有个全面的认识和把握;在撰写综述的过程中,确实应该强调近期的文献(特别是最近5 ~10 年的),因为这样才能更有力地说明某个领域的现状,但是对于一些开创性的、影响深远的经典文献,还是应该包括在综述之中的。

讲到这里,如果你还在为怎么把所有东西塞进 25 页纸里而犯愁,那么你很可能犯了第二个错误:你并没有明确地限定综述的范围。虽然不必把每一篇读过的文章都拿来引用,但你却需要一个判断标准来决定取舍。切忌用“我喜欢这个研究”或是“我没看懂这篇文章的数据”作为取舍的标准,因为这样的标准过于主观。可以用被试的某种属性(如:年龄、民族、性别或是其他与研究相关的特性)作为限定文献范围的标准,也可以用设计的类型或质量、自变量或因变量来限定文献的范围。越早提出这个限定文献的标准,对你的写作就越有好处,它能有效地减小主观情感(例如,“这个研究其实并不支持我的观点。”)对你的选择造成的影响(Cooper, 1998)。另外,要记得把限定标准也写进文献综述的引言中。

举个例子,有一个项目旨在教会儿童如何有效处理来自同辈的挑衅,而 Julie 准备做一个关于评价这个项目的研究。于是她收集了大量有关同辈相互排斥的、社会技能培养的以及一些与行为改变有关的文献资料。经过一番思考和讨论,她意识到其实只有少数有关社会技能的研究关注了同辈排斥这个问题,而大多数社会技能的文献都与她的研究无关。如果她在自己的文献综述中逐一介绍这些文献,那么 25 页的综述将大部分被它们所占据,而其他真正有用的文献都只能是附带地提一下了。最终,她选择对有关社会技能的文献仅做简要概述,而不是主次不分地逐一描述。可见,通过这种对主题范围的限定,可以使我们更好地组织我们的综述。

另一名学生 Lisa 也遇了这样的问题。她发现自我暴露疗法对青少年受训者参与性的影响是一个很有趣的题目，于是她通过检索得到了大量的文献，其中包括：成年人在心理治疗过程中的自我暴露、成年人和青少年在人际交往中的自我暴露、不同的心理治疗对青少年的效果以及人际交往中自我暴露效应的理论阐述等。可是，这些都与 Lisa 想要研究的题目不完全相同啊，她的文献综述中到底应该包含哪些研究呢？Lisa 最终决定深入地评述两类主题的相关文献，首先是成年人在心理治疗过程中的自我暴露的相关文献，因为它们既能帮助她证明她所设置的自变量的合理性，又能突出研究方法的创新性；其次是自我暴露在青少年人际关系中的作用的相关文献，因为它们能够帮助她解释为什么自我暴露在心理治疗过程中扮演着重要的角色。而对于其他的文献，仅在有助于描述上面提到的两类主题时才作简要地概述。

编排文献

无论是拟定写作提纲还是阅读文献，都需要思考如何对文献加以组织和整合。现在你已经写好了写作提纲，该是往提纲里添加细节的时候了，你可以用不同的方法把文献分成易于管理并且连贯的小组。以下我们列出了一些最常用的方法：

1. 按照自变量分组。如果研究的是关于提高组织效率的不同策略，那么就把文章按照策略的类型分组，这样会形成一个容易上手的框架。

2. 按照因变量分组。比如，研究酗酒青年人的特点时，可以将文献分为探讨他们人格特点的研究、对他们酗酒方式的研究以及对他们人际交往技能的研究。此外，还可以按照评估方法（如自我报告法、他人评价法或直接观察法等）对研究结果进行分类。

3. 按照实验设计的类型分组。一般来说，这种方法是将文献按照其研究设计控制的强弱程度分为不同的组，例如：无对照组的设计和对照组设计、相关设计和实验设计、横断研究设计和纵向研究设计等。Paul（1969）介绍系统脱敏研究的经典综述中就用到了这种分组策略。

4. 依据理论前提组织文献。当你的研究是在检验对于一种现象的争论，或者不同的理论线索都有助于你对研究问题的系统阐述时，这种组织策略就能派上用场了。比如说，Wolfe（1986）曾对家庭中父亲与行动异常儿童（或无行动异常儿童）的相互作用进行了研究，她以社会学习理论的视角，围绕着父亲的哪些行为会影响儿童破坏行为的产生和持续，对文献进行了组织。（参见专栏 6.4）

这些策略只是众多组织文献方法中比较有代表性的几种。选择一个能够突出文献重要方面的组织框架，这种框架最好能够突出在你研究中希望改进的因素。比方说，假如你的研究相比过去的研究在方法上有了改进，那就将获得因变量的方法作为文献分类的策略，这样读者可以更清楚地看到过去的研究方法存在的缺陷。在这个过程中记得采用漏斗式的写法：从一般到具体地去写。

大量使用副标题也是一个不错的办法，这表明你对综述的主要部分已经进行了深入的思考，而且这样做也能帮助读者理解文章安排的逻辑。此外，还有一个很好的方法就是随着文章的推进对之前的内容进行总结，这在一个或几个特别长或者特别复杂的部分之后这样做是特别有用的。记住，你的目标是引导读者进行阅读，并

对为什么你的研究将成为这个领域中的关键一步进行有力的论证。

专栏6.4

依据理论前提对文献进行分类

注：转自 *Paternal and Marital Factors Related to Child Conduct Problems*, by V. V. Wolfe, 1986, Doctoral dissertation, West Virginia University. 使用经过作者授权。

开始写作

读到这里，相信你已经做好动笔写作的准备了。那么接下来，请整理一下笔记并拿出提纲，开始撰写综述每个独立的部分。

在这个时候作者往往会有一种幻想：我所提交的综述将会是完美无瑕的，在这篇文章中所有以往的相关研究必将被一一列举、所有见解都将会是卓尔不群的！这种期望影响了许多学生，哪怕是那些经常不费吹灰之力就能写出文献综述以应付课堂作业的学生，也免不了落入此俗套。

可是事实并不像我们想象得那样完美。大多数学生在他们最终完成文献综述之前都要几易其稿。不管你的文章在撰写过程中多么的字斟句酌，答辩委员会的专家都会提出他们的建议，并希望你能够修改初稿。你的草稿不可能是完美的，所以请你容忍这种瑕疵，并给自己的文章留下能够被进一步完善的余地。请你记住，犯错误也是一种积累学习经验的过程，如果你的初稿就是十全十美的，试问你在撰写文章的过程中又会有什么收获呢？

一些学生在动笔开始写作时还可能会错误地认为他们必须严格地遵守写作次序：从引言开始，按部就班地依次往下写。这种教条的错误可能会导致他们枯坐在稿纸或者电脑屏幕前几个小时，仅仅是为文章开头的一句话怎么写而绞尽脑汁。或许有这样一种更好的思路，你不妨在完成了大体的文献回顾后再回头去撰写引言部

分的内容。这样做的好处是,不仅能够对整个文献回顾的结构有一个很好的认识,而且也会在引言部分对文献回顾部分有一个良好的概括。临床心理学家 Bolker (1998)就曾经给过作者们一个忠告:对于大多数人来说写作是一个棘手的过程,在此过程中,哪怕我们只是在将一些句子混乱地罗列在一起,抑或我们只是在生产垃圾文章,我们也依然要坚持不懈地完成整个过程,而不要因为受制于写作顺序的教条,在写作的一开始就止步了。

如果你觉得从引言开始写是一件困难的事情,那就找一个你觉得容易的部分下手。如果觉得一下子写出一个段落对你来说也不是件轻松的事情,那就先写出这一个段落你要表达的中心思想,然后围绕着这个中心思想来建构整个段落。如果你没有想好用什么样的过渡方式将几项研究串联在一起,那么就先将研究的内容交代清楚,然后再构思它们之间该怎么过渡。你觉得困难的部分不会随着时光的流逝而变得简单起来,所以无论如何要找到一个你能够写出来的部分立刻动手。请一定要记住:修改已经写出来的文字远比你最初写这些文字时要容易,所以至少先动手写出一些文字来,这样我们才有调整顺序、整合和修改润色的对象。

此外,我们在第 3 章提到的一些办法在这里也是适用的。如果你不知道从哪里着手,不妨先设置一个小目标。目标可以分为两种类型:一种指向结果,一种指向过程。指向结果的目标对你需要完成什么任务(例如“写出论文的一页”或是“总结并评价一篇重要的文献”)作出规定;指向过程的目标也是对你需要做什么做出规定,但是这种目标并不关注最终的结果(例如“花两个小时的时间撰写学位论文”)。我们希望你能够为自己设置一个指向结果的目标,因为你可以发现即使坐在电脑前什么都不做也能完成指向过程的目标!一些小的指向结果的目标可以帮助你清晰地了解你的工作进展到哪一步了,并且也可以提高你的工作效率。

当你实现了一个小的目标时可以给予自己小小的奖励。还记得我们第 3 章提到的普雷麦克原则吗?任何发生概率高于你写论文这件事的行为都可以被视作是小小的奖励,对于多数作者来说这个时候一定在感叹似乎任何事情都可以作为奖励吧。其实在我们撰写这本书的过程中,我就经常用这个原则来奖励自己,我时常对自己说:“等我写出来 3 页就奖励自己一杯咖啡喝”“完成这个表格我就可以读一会杂志了”,或是“完成这一章的编辑我就可以去给朋友打电话了”。请注意我们所说的奖励不需要耗费大量金钱或时间,能起到激励你进一步去完成接下来的目标的作用就可以了。

撰写引言

用一个简要的引言作为文献综述的开头。在这里需要做的事情是:①简洁地介绍和描述全文主题;②解释核心概念;③交代你这篇文献综述的范围和大体结构。在一篇 25 页左右的文献综述中,一段理想的引言大约占 3 页左右。

你可以将交代文献综述的范围作为引言部分的最后一段,这一段将作为后文的先行组织者(advance organizer)——对文献的覆盖范围和接下来读者将会读到什么加以介绍。除此之外,如果为了控制文献综述的长度而有意忽略了一些研究,那么请告知读者你选择或忽略研究的标准。例如,一类与你的主题相关的研究后来被发现存在特殊的方法学问题,可能你觉得在文章中没有空间去交代这些研究,那么要

向你的读者说明这件事情。综上所述,引言部分的最后一段可以这样写:

> 接下来此文将对新近文献关于________问题的看法进行评论性的回顾。这个领域的早期研究(例如________)通常认为________(2004)。但是,就像________(2004)和________(2005)所指出的那样,这些研究都没有对______(2004)和________(2004)进行区别,所以它们的结论是有待商榷的。新近的研究已经对上述缺陷加以改进,对此在下文中会加以详细的介绍。在文章的开头,讨论了________的问题。这篇文献综述对现有文献进行了总结和评论,此外还对具体的研究问题和研究假设进行了讨论,并将在后文中对提出的问题加以检验。

撰写文献综述的主体部分

如果说引言的功能是概括全文,那么在引言之后的主体部分则起到了回顾和整合以往文献的作用。对主体部分的编排要遵循你在拟定写作提纲时划分文献类型所用的规则(依据自变量、因变量或实验设计类型等),将有关联的研究归为一组,再按照文献发表时间的先后顺序逐一介绍。在介绍每项研究时要使用并行结构,即采用相同的方式描述不同研究的各个方面,正像 Bem(1995)所说:"重复和使用并行结构是使得文章清楚明了的最有效的方法,在这一部分不要加入你自己的创造,清楚明了是最重要的。"(p. 174)只要将每项研究的被试、自变量、因变量、实验设计、实验结果和其他值得关注的细节(如方法的问题、被试的种族和性别等)交代给读者就好。在这一部分的最后,请对所有的文献加以总结和评价。

按照上面提到的方法,主体部分的写作似乎轻而易举地就完成了,但是当实际操作起来时,你会发现自己可能会将这一部分写得像流水账一样苍白。这里就要用到诸如过渡语句或是重申主题这样的技巧。另外,简要的比较实验方法和实验结果也可以有助于将不同的研究组织到一起,而且这样作还能突出它们的异同。Inderbitzen-Pisaruk 和 Foster(1990)的文献综述就用到了上述表达方式,接下来我们将节选他们这篇文章的片段,并删去主要内容而只留下该片段中的过渡语句供大家参考。

> Kuhlen 和 Bretsch……(对研究进行描述)……。
> 在一项研究中认为应将行为归为……必然的结果……。
> 但是不幸的是,上述的两项研究都过于陈旧,因此……。
> 除此之外,它们都……。这使得是否……变得不确定……
> 这种不确定性被越来越多最近的研究所关注,有趣的是,尽管青少年文化在过去的四十年发生了巨大的变化,可是新近研究的结果还是与上文中提到的那些研究大同小异,例如……(Inderbitzen-Pisaruk & Foster,1990,pp. 426-427)

另一种组织和编排主体部分文献的方法是:将那些不具备卓越学术价值的或是使用了近似方法的研究放在一起,对它们只做简要的回顾,然后将更多的精力放在那些有影响力的或是意义重大的研究上。使用这种编排方法,你可以先在几个段落中根据结果的优缺点对那些普通的研究加以描述,而后再详细地阐述那些有创见的研究。

第三种方法是依据研究的结果编排文章的主体部分,这种方法的特点是:它所包含的对个别研究的描述相对较少,相反,作者会使用其他研究的结果作为文献回顾中

观点的佐证。使用这种编排方式的文献回顾是最难写的，因为你必须十分有逻辑地将自己的观点引出，并且在使用文献（赞成的或反对的）支持自己的观点时必须要做到不偏不倚。当你的研究想要介绍两种对立的理论解释，抑或你所选择的变量更多地依据理论而非实证结果时，这种需要做多方面考虑的方法可能是最适合的。

上述3种方法都是对文献定性的分析和整合，我们要介绍的第四种编排方式则是一种定量的方法，这种方法按照不同研究的特性为研究评分，然后再对其进行元分析（meta-analysis）。在元分析中，那些具备相同特质的研究结果被按照一定的统计规则归为一类，并被用来估计这类研究整体中各个特性的情况，以获得普遍性、概括性的结论。尽管元分析这种方法在综述性的学术期刊中是很常见的，但是在学位论文中还是很少被用到（可能是因为做一个元分析简直就像做了一个实证研究一样，将会耗费研究者大量的时间和精力）。所以在这里我们不再对元分析进行详细的介绍，Cooper（1998），Hunter 和 Schmidt（2004），以及 Rosenthal（1991）对怎样进行元分析做了简明扼要的介绍；Lipsey 和 Wilson（2000）为元分析的初学者提供了一个指南，建议每个新手都看一看；此外，Rosenthal（1995）将告诉你如何运用元分析方法撰写文献综述，如果你对元分析感兴趣，可参看上述资料。

对文献进行整合和评论性的分析

教授们经常会抱怨他们的学生不会对文献进行分析和整合（Association for Support of Graduate Students，1995），文献综述的新手通常会将文章写得冗长且缺少必要的评论性分析。请不要仅仅是将读到的东西写出来，还要对文献进行整合。你在前人的研究结果中发现了怎样的模式？这些结果一致吗？如果不一致，为什么呢？想一想这些不一致潜在的解释方法：是方法学上的问题？实验设计的问题？不同实验中样本存在差异？还是自变量或因变量控制或设置的问题？当读者看到文献综述的最后，应该能够对你研究领域中的研究结果和研究方法的模式有清晰的认识。

除此之外，你还要对文献进行评价性的分析。哪些研究是杰出的，为什么？想一想它们存在哪些优点和不足？这个时候你可以用到专栏6.1提供的核查清单。即使发表了的文章也不代表它们就不存在任何方法学的问题，你需要帮助读者看到在这个领域未来的研究中应该关注到哪些方法学的问题。如果你在之前的阅读中，坚持将你认为文献中存在问题的部分记录了下来，这个时候就可以找出笔记，参照这些问题想想你要在文章中进行哪些评论性的分析。

在评论其他研究者的工作时要保持专业的素养，特别是当你的论文引用了论文答辩委员会主席的研究时，不要将其贬斥得一文不值；此外，最好也不要在评价与自己观点相反的研究时过于激愤，人身攻击是绝对不被允许的。请保持冷静和公正，时刻记得任何的研究在某些方面都不可避免地会存在缺点，同时，在另外一些方面，它们也一定有值得你学习的地方。

对文献的整合和分析能够帮助你为自己的研究做好准备，当读者看完你的文献综述也能够清楚地认识到研究问题的基本逻辑以及你采取某种研究方法的原因。因此，对文献的整合能够将尚未解决的问题显露出来，而评论性的分析不仅可以揭示过去研究方法上存在的问题，还能告诉读者以往方法上存在的问题会在你的研究中得到怎样的改进。想一想在这个过程中，你就好像是一个优秀的律师：试图去使

陪审团(读者)意识到某条证据与你受理的案件的关系是多么的密切。

整合和分析文献有许多种方法。首先,你可以使用过渡句或短语去帮助读者发现他们正在阅读的文本中包含的模式。请阅读下面的段落,看看它是如何将一类研究编排在一起,又是如何突出它们与将要被讨论的研究的主要异同。

> 尽管以往大多数研究都采用了相关设计,但是 Smith 和 Jones(2001)却采用一种实验设计的方法研究了________问题。他们得到的结果与以往使用相关设计的研究没有显著差异。

另外,可以使用比较的或评价性的表达方式,当然要注意一些评价性的评论只适用于一些特定的研究。下面的这个例子或许能够对你在描述研究时有所启发。

> 尽管作者的研究结果与其他研究者的结果存在差异,但是他们所抽取的样本容量过小(只有 8 个)可能会严重制约他们结果显著性的统计效力。

还可以使用下面这种评价方式,它比较适合作为对一类研究的评价(注意:不是对整个研究领域的评价),你可以将它们放在对某一类文献回顾的最后。

> 多数这类研究存在相似的优点和不足。(加以详细描述。)尽管它们存在方法上的问题,但是多数实验得到了相似的结论。(加以详尽说明。)

你可以在最后对全部文献进行总结和评价,这个部分可以列出尽管现在未被研究者关注、但在未来研究中可能具有价值的题目,也可以对过往所有文献的优缺点加以总结。不要把那些在前面说过的内容再度拿到这里来充当总结的内容,而要利用这部分梳理清楚文献回顾的思路。更重要的是,你还应该将这部分视作是后面问题的提出和研究假设部分的逻辑导线。

如果你还需要一些实例,请仔细阅读发表文献综述的那些学术期刊,例如《心理学公报》(*Psychological Bulletin*)、《临床心理学评述》(*Clinical Psychology Review*)或是《发展心理学评述》(*Developmental Review*)等。除此之外,一些图书中的章节也会提供文献综述的合集。你可以从中分析各种文献综述的结构和组织方式,探究不同作者是如何整合和分析文献的,在这个过程中你会获得如何撰写综述的信息。1995 年的《心理学公报》(*Psychological Bulletin*)曾以专题论文集的形式发表了一系列对撰写文献综述有益的文章,其中 Bem(1995)还对学位论文与专业期刊中文献综述的差异发表了有趣的见解。

介绍你的研究和假设

如果你所撰写的文献综述遵循的是一章模式,那么问题的提出、研究的问题以及研究假设通常要写在总结和评论部分后面。如果你遵循的是两章模式,那么上述的几项通常是要写在引言之后和文献回顾之前的。在文献回顾之前交代研究主题可以帮助你确定文献回顾的范围,此外文献回顾部分也可以为你之前提到的研究主题提供理论依据,并且能帮助读者了解为什么你的这项研究将会对这一领域有新的贡献。在撰写这一部分的时候请记住读者在读到你的文献回顾之前,是要先读到你的研究主题和假设的,所以要向读者提供他们可能暂时不知道的背景知识,使得读者能够清楚地理解这一部分的内容。

无论选择哪种写作模式，都要在问题的提出部分交代你这项研究的理论依据。此外还应该简要地介绍一下你选用的样本、实验设计、自变量和因变量，这些都将有助于读者更好地理解你将在后文中写到的研究问题和研究假设。

第4章中已经介绍过如何陈述研究问题和导出研究假设，现在你可以依据我们的那些建议进行这部分的写作。如果研究假设的理论依据不是十分的清晰，你需要简要地告诉读者你做出这个预测的原因是什么，比如说你可以这样写："本研究假设是依据 Smith(2004)的研究结果提出的，Smith 研究发现________。"就像第4章提到过的那样，你可以将研究假设和主要的研究问题联系起来。当然难以确定研究假设的情况也是存在的(可能你所研究的是一个全新的领域或与前人的研究结果相互矛盾)，在这种情况下如果真的无法确定自己倾向于哪种研究观点或理论，则可以采取以下3种处理方法：第一，可以直接向读者说明你这项研究的假设难以确定具体的方向；第二，可以将研究问题再细分为研究关注的主要问题(提出这一部分的假设)以及研究关注的次要问题(这一部分不做假设)；第三，还可以将难以确定研究假设的问题排除在你的研究之外。尽管我们推荐前两种处理方法，但不同的论文答辩委员会对此有不同的观点，所以在这个问题的处理方法上你还是要请教一下论文答辩委员会的专家。

撰写附录

我们上文中已经介绍了多数学校对学位论文中分支部分的要求。除此之外，一些学校还会要求学生提供一些附加的材料，如对术语的解释、阐述你提到的研究中存在的局限或是讨论研究背后的理论依据等。所以，请检查一下你是否清楚学校要求提供哪些附加材料。

避免抄袭

我们在前文中提到过抄袭的问题，到底什么是抄袭呢？抄袭是指"擅自将他人的思想或语言说成是自己的观点：未在参考文献中标记就使用他人的研究成果的行为"(Merriam-Webster，n. d.)。由上述定义我们可以知道未标明引用而使用他人说过的话就是赤裸裸的抄袭行为，此外，将他人的观点用你自己的语言改写、照搬他人文章中段落的编排方式或是使用他人组织主题的方法也都是抄袭行为。想避免这些抄袭行为，就需要你使用自己的语言并且按照自己的方式编排和组织你的观点，当需要借用或引用他人的观点或组织框架时一定要注明，并且只允许自己进行直接引用。如果你对抄袭这个话题还存在什么疑问，可以去询问你的同学或老师。

修改初稿

请做好修改论文初稿的心理准备。当你再次阅读自己写的初稿时，可能会发现自己对某些问题又有了新的想法，也可能会发现论文在逻辑上存在的某些漏洞。此外，你的导师(或是你的同学)也会对你的论文提出修改意见，请虚心地对待他们的批评和建议，因为这些都将是对你有益处的。一些章节可能你自己感觉写得十分完美，但是读者读起来却可能有种云山雾罩的感觉，所以还是请牢记你是在为读者写作，而不仅仅是在自娱自乐。

你的目标是写出一篇可以通过论文答辩委员会审查的论文，所以请把这个过程当成是一次提高写作能力、思维能力和掌握某一领域文献能力的机会，让你的论文成为一流的作品而不是被别人抨击的对象。无论你今后选择怎样的道路，在撰写学位论文这个过程中磨炼出来的写作技能和思维方式都将使你受益终身。

将你在文献中获得的信息整合在一起，使其成为连贯且有侧重的文献综述，形成这种能力远比要你完成一篇文章的本身更具价值。因为它既能够使你对研究主题有敏锐的思考；同时也能帮助你对文献中出现的概念和方法有清晰的认识；此外，还能引导你确定在研究中将要运用怎样的方法和程序，由此可见撰写文献综述的过程和确定研究方法这两者是密切相关的。研究方法是指你将选择怎样的方式进行研究，我们会在接下来的两章中对如何撰写研究方法部分做详细的介绍。

补充资源

Galvan, J. L. (2004). *Writing literature reviews: A guide for students of the social and behavioral science* (2nd ed.). Los Angeles: Pyrczak.

Heppner, P. P., & Hepperner, M. J. (2004). *Writing and publishing your thesis, dissertation, and research*. Belmont, CA: Thompson-Brooks/Cole.【与我们这本书类似，该书中提到了更多有关写作的实例。】

有关抄袭

Plagiarism: *What it is and how to avoid it*. Retrieved February 4, 2005, from http://www.indiana.Edu/~wts/pamphlets.shtml.【这是一本由印第安纳大学（布卢明顿）写作指导中心组织编写的简明手册，它将通过实例告诉你哪些行为算是抄袭而哪些行为不是。】

Sandler, R. B. (2000). *Plagiarism in colleges in the USA*. Retrieved February 2, 2005, from http://www.apa.org.【作者是位律师，这篇文章从法律的角度诠释了抄袭。】

有关写作过程

Bolker, J. (1998). *Writing your dissertation in 15 minutes a day*. New York: Henry Holt.【概述了我们这本书有关如何全面高效地撰写特定领域的学位论文的想法。】

Vernoff, J. (2001). Writing. In K. E. Rudestam & R. R. Newton (Eds.), *Surviving your dissertation* (2nd ed.). Thousand Oaks, CA: Sage.【对导致写作问题的原因有其独特的见解，并提供了一些学位论文写作的建议。】

思考与练习

文献回顾

☐ 查找相关文献

—确定主要研究者和重要的期刊

—使用参考文献索引

—运用计算机检索文献

—索要文献

—查看其他学科的相关文献

思考与练习

—浏览核心期刊目录
—利用参考文献列表
—使用原始资源
—避免使用大众媒体

□ 批判性阅读
—确定主题
—找出每篇论文的优点和不足
—总结该领域研究的优点和不足
—整理影印本和笔记

□ 写作准备
—了解文章长度和格式的规范
—写一个初步的提纲
· 限定文献综述的范围
· 组织和编排文献
· 给拟撰写的每个部分分配页数

□ 撰写文献综述
—撰写引言
—撰写主体部分
· 使用过渡性语句和整合性短语
· 对文献进行整合和评论性的分析

□ 介绍你的研究和假设
□ 避免抄袭
□ 修改初稿

7

研究方法和伦理道德

本章我们将具体阐述研究方法的组成要素，然后讨论心理学研究的伦理原则，并依据这些原则来判断研究设计和研究过程的好坏。研究过程包括获得被试的知情同意、数据保密以及向被试告知相关信息等。此外，本章也涵盖了如何与机构审查委员会（IRBs）就保护人类和动物被试进行交涉。

需要注意的是，本章不是专门讲授研究设计、步骤和伦理的。本章不会告诉你你的实验应该使用哪种设计（相关设计还是区组设计），也不会告诉你是否已经全面控制额外变量。如果你对这些基本的研究方法很生疏，可以参考本章的附录。

了解方法部分的组成

《Publication Manual of the American Psychological Association》(APA,2001)描述了方法部分的3个基本要素:被试或主试、仪器和程序。一些研究严格按照这种方式阐述,但多数研究也包含其他的要素。下面,我们介绍一个完整的方法部分中所包含的全部要素,但请注意,这些要素及其编排的顺序并不适用于所有研究。在对方法部分的各要素进行详细说明之前,我们先要简略说一下研究设计的问题。

研究设计

上述手册并没有明确地把研究设计看成是方法部分的组成要素,这是因为研究设计通常以段落的形式放在方法部分之前,或者在前言中的"问题陈述"部分里来简要说明。由此可见,研究设计的位置是灵活的,我们应该把它放在对研究最有意义的地方。我们建议把研究设计放在方法部分之前,这是因为研究设计会影响方法部分的所有要素,把它放在前面会有助于读者更好地理解方法部分。但是,当设计很复杂、需要进一步解释的时候,要把研究设计放在方法部分之中。

研究设计里都需要说明什么呢? 首先,要明确自变量类型:是被试内变量还是被试间变量。如果研究的是同一组被试在不同时间或不同情境下的变化,那么就是被试内变量;如果考察的是在同一时间点上不同被试间的差异,那么就是被试间变量。其次,阐述研究设计的一般类别,是实验的、准实验的、相关的、纵向的还是个案的。相关内容可参考 Shadish,Cook 和 Campbell (2002)。

弄清楚研究的基本性质之后,就可以详细说明你的设计了。

在这一部分不要混淆研究设计和统计,"采用单因素方差分析检验各组间的差别……"是对统计检验的描述,而不是研究设计。类似的表达应该放在分析(在结果部分)里,见第5章。试举两个典型的例子:"把儿童随机分配到高、低奖赏两组和高、低数学任务难度两组,这是一个2×2的因素设计。""这是一个多基线设计,第一个/组家庭首先实施积极的家庭规则,然后第二个/组家庭开始实施,最后是第三组家庭。"

如果进行多种测量,需要安排测量的顺序(如平衡),并在研究设计部分进行说明。如果需要对被试的某些特征(如性别、IQ或问题严重性)进行匹配,需详细说明其过程。进行平衡和匹配是必要的,要了解这些操作的原因。

图示有助于描述复杂和难于理解的设计。沙迪什等(Shadish et al. ,2002)提供了一个很好的关于被试间比较的例子。海斯,巴洛和纳尔逊-格雷(Hayes,Barlow & Nelson-Gray,1999)提供了许多关于被试内比较的例子。图7.1是索贝尔,博加德斯,舒勒,利奥和索贝尔(Sobell,Bogardis,Schuller,Leo & Sobell,1989)的一项研究中的图示,虽然很复杂,但能够提供很多信息。

被试

在被试部分需要交代3个问题:①谁将参加实验? ②有多少人参加? ③如何选择被试? 如果被试不是随机分配,需要根据被试的某些特征进行分组(例如根据年龄、诊断分类或是根据特定任务的表现来分组,注意:此分组本身就是一个自变

量),并且要说明分类的过程。

不同的研究中,被试部分的描述有很大不同。被试是一个更大群体的子集,被试的选择要以能够代表这个群体为标准。例如,如果被试是人类,最好提供一些人口特征的信息,如性别、年龄(平均年龄和年龄范围)、种族、代表的地区、社会经济地位、来源(如大学本科的班级或者心理健康诊所)、参加的条件(如自愿、报酬或学分)。也可以根据研究性质提供一些其他的信息,如智力机能(在一个标准 IQ 测验上的平均成绩及其范围)、在选择变量上的得分(如抑郁、焦虑、超重比例)、残疾状况、诊断和性取向等。如果某个被试特征是自变量,需要把它描述清楚,如怎样判断被试是否有此特征。另外,需要说清楚具有哪些特征的人不能参加实验(即排除标准及其操作过程)。

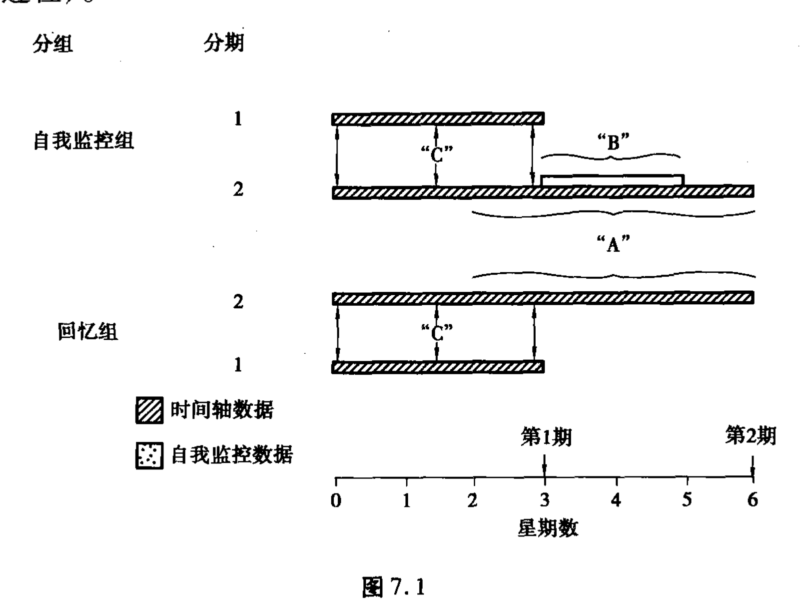


图 7.1

实验设计示意图,两组被试的时间轴和自我监控饮酒数据,每一期的时间安排,主要统计分析(用图中的字母表示)中的数据集。A 对照是被试间分析,采用第 2 期的时间轴数据来检验自我监控的反应。B 对照是被试内分析,分析自我监控组的第 2 期自我监控和时间轴数据;比较同一组被试在同一时期内的时间轴与自我监控数据。C 对照比较每个被试在第 2 期的时间轴数据与第 1 期中相同时间间隔内的数据(通常是 21 天)。转载自"Is Self-Monitoring of Alcohol Consumption Reactive?" M. B. Sobell, J. Bogardis, R. Schuller, G. l. Leo, and L. C. Sobell, 1989, *Behavioral Assessment*, 77, p. 451, 版权 1989, 得到 Elsevier 的许可。

初次做研究的人有时不知道研究中谁是被试,特别是当不同的被试组在不同时间参加实验的时候。例如,在建立评估花销问题的尺度时,我们的一个学生请来消费信贷专家,让专家判断人们在消费问题情境中的各种反应的适当性。她用这些判断来衡量个体的反应,进而建立初步规范。因为研究的首要目的是建立一个可靠有效的用来识别花销问题的工具,能够提供与信效度有关的数据的个体就是被试,而提供适当等级的个体是评判员。当然她也需要在方法部分详细说明一下评判员的特征。总体来说,在被试部分所描述的个体是那些为检验研究中特定假设提供数据的人,而那些生成工具或收集数据的人(评分者或观察者)不是被试。

说明取样策略及招募被试的过程。是在当地报纸上登广告?是在校园里张贴

广告？还是从出生登记册里选择？说明具体做法，并把招募被试所拟用文档的复件放在附录里。

如果研究中采用的是非人类被试，也要说明上述类似的问题。说明研究中动物的数量、种类、物种和品种编号。此外，有关被试来源的信息也很有用（如提供者、库存名称）。最后说明它们的性别、年龄、体重及其一般条件。当然，关于它们的历史、护理及处理的重要细节也要进行描述（APA，2001）。

说明潜在被试的数量及实际参加实验的人数，包括那些因为与标准不符而被排除的人数。被试的代表性对于研究的外部效度是至关重要的。因此，除了在最终稿中报告上述信息，最好把你所能知道的有关被试与非被试之间差别的信息也报告出来。如果有一些被试中途退出实验，还需报告这类人所占的百分数，并说明离开的原因。

选择被试时的经验法则是“代表性”原则，即考虑被试所代表的更大群体。清楚了这一点，就可以设计程序来招募被试。可以收集相关信息来看看被试与所代表群体特征的相近性，以此来考查样本的代表性。有时你可能不想把研究结果推广到某些群体中，例如，一个关于低年级儿童的研究并不关心其研究结果能否推广到高年级儿童中。为了有效地利用你的时间和其他资源，可以尽量采用能够筛选出不适合研究的个体的程序和手段。

研究设计的特定性质也能够决定被试是否能够参加研究。总体来说，同质性群体比异质性群体在测量中分数变化相对较少。这是因为在异质性群体中更有可能存在一些与研究不相干的，但与测量分数相关的被试特征。一方面，在采用 r 及其变式（如多元回归和因素分析）的纯粹相关设计中，自变量与因变量分数的变异性很重要。这是因为相关变量的限制范围约束了相关能够得到的数值（Anastasi & Urbina，1997）。

另一方面，在采用方差分析和相关途径来比较离散群体间的差别时，没有得到控制的、与研究不相干的及变化的被试特征都会产生误差方差。如果没有控制额外变量，那么就需要更有力的自变量来得到显著的结果。在这些情况下，限制被试群体是很有必要的。找出那些对结果不太重要的（如社会经济地位），但与主要因变量相关（如智力）的被试特征。根据这些特征可把被试限制在某个子群体中（如在一个相对狭窄的社会经济范围内的所有个体）。

在区组设计中可通过两种方式来解决被试特征中的变异。可以把一个或多个被试特征设为自变量。例如，在研究有关学前阅读测验分数上的观看“芝麻街”效应（the effects of watching“Sesame Street”）时，如果关注社会经济地位，那么可以把社会经济地位设为自变量。当然，这会增加被试的数量。另外，可通过统计来解决：收集完数据后，看看社会经济地位是否与因变量相关，如果相关，那么在随后的分析中可把社会经济地位设为协变量。

收集被试信息是很必要的，部分信息是关于排除掉的个体，部分信息是关于被试的适当描述。特别要关注那些能够对你的预期结果提供另类解释，或者无意中会与自变量相混淆的被试信息。可以通过查找文献来找出潜在的混淆变量：哪些被试变量与因变量相关？测量它们并随后进行分析，看看它们能不能被排除掉，以此作为对结果的解释。

还需要说明一下这些被试特征是如何测量的。如果想采用一个人口数据表和两个问卷调查来收集被试信息。首先，确保它们没有版权，并把它们放在附录里以供读者参考。同时在被试部分或随后的测量部分说明一下这些工具的心理测量属

性(信度、效度)。

最后谈一个普遍问题——"我需要多少被试呢?"一种途径是通过阅读文献来看看标准是什么。这很容易,且适用于个案设计,但对于其他研究类型来说,并不是一个好的方法。通过翻阅那些最有威望的专业杂志你会发现,多数区组设计的样本量太小以至于得不到足够的统计检验力来检测出那些最强大的效应(Rossi,1990)。

那么我们应该用什么方法计算被试数?在回答这个问题时需要考虑两件事情,一是所采用的统计方法。为了得出稳定的结果,一些分析方法对于估计被试数量与变量之间的最小比率有一些经验法则。采用多元回归、因素分析、多元方差分析(MANOVA)、判别函数分析,或者相关统计时,从统计经验法则的角度来计算被试人数以达到最低要求的方法是:数一下因变量将会产生的分数的个数,把它们排序形成子集,每个子集需要包含在每个特定分析中的分数的名字。例如,采用一个包含3个不同因变量的多元方差分析来评估侵略性,这是第一个子集,第二个子集包含5个人口学变量,用于再一次的多元方差分析,等等。数一下在每个子集内的因变量的数目,然后用最大的数目(在上面的例子中是5)乘以在7~20之间的被试数目(保守的统计学家会在每个变量下用20个以上的被试,但有些人会用7个),得到的结果就是所需要的最少被试数目。

统计专家有时批评经验法则过于简单(如,MacCallum,Widamon,Zhang & Hong,1999)。第二种方法也许能更好地计算被试人数,这种方法需要把统计检验力考虑在内。检验力与计算被试人数的经验法则无关,它是指统计方法能够检测到自变量效应的可能性。检验力是3个事物的函数:①样本大小;②自变量效应的大小(即自变量对因变量贡献了多少差异);③α 水平。通过检验力分析能够让你在检测小、中、大效应的时候估算出各需要多少被试。设定被试数目和 α 水平(通常 $p<0.05$)后,所期待的自变量效应越大,用于检测它的检验力越小(即需要"更少被试")。

计算区组设计的检验力,首先需要计算出想要检验到的最小效应。检测到微小的效应意味着需要成千上万的被试(见 Cohen,1992)。"在这个领域里有意义的最小效应是多大呢?"在临床上需要达到中等效应——小于中等效应可能达不到实践上的显著。可见,可被接受的效应大小取决于具体研究领域。在一些很多因素会影响到因变量的领域,如药物的某些方面,记录小效应是很重要的。

在进行检验力分析之前,需要掌握一些信息。第一,需要选择检验力水平:我们建议选择0.80。第二,需要知道所采用的分析方法(是方差分析、回归分析,还是相关分析)。此外,需要明确分析中的哪些方面能够看出数据是否支持预测。例如,仅仅知道要做2×2方差分析是不够的,还需要知道假设能否预测主效应或交互作用。同样,做回归时,需要弄清楚是否所有的变量预测同一个结果(即总体 R^2 显著),或者是否每个变量都会增加显著的独特方差(即 β 显著)。用统计参数来评估结果是否显著,并根据统计参数来判断检验力。

第三,需要根据对可能收集到的数据的猜测来估计实际人数,进而计算检验力。例如,如果采用参数统计,就需要提供平均数、标准差及 *ds*(效应大小值)。如果采用相关统计,则需要提供预期的相关值。你可能会问,我怎么能在做研究之前就得到这些数值呢?你当然不能,至少不能精确人数。但是,可以根据先前收集到的数据,或者根据采用相同因变量和相关自变量的已发表研究来进行估计。然后,就可以把这些值代入公式中来计算所需要的被试人数。如果你想对统计检验效力有进一步的了解,请查阅:科恩(Cohen,1988)及凯佩尔和威肯斯(Keppel & Wickens,

2004)关于如何计算检验力的文章;汤姆斯和克雷布斯(Thomas & Krebs,1997)对用来计算检验力的计算机统计程序所做的回顾。此外,一个用来进行简单检验力分析的互动练习网站是 http://calculators. stat. ucla. edu/powercalc/ 。这个网站的软件是免费的,只需按照要求输入一些数值,它就会立刻为你进行分析。需要事先提醒注意的是:网站中一些分析的名称与心理学的惯用术语是有所不同的。

另外,科恩(Cohen,1992)简洁通俗地讨论了检验力分析,并提供了一个用于估计被试人数的表格,这个表格告诉我们在 $p<0.01$ 及 $p<0.05$ 水平下观察各类效应大小所需要的被试人数。它涵盖了 t 检验,卡方检验,方差分析和多元回归统计。另一个比较容易使用的资源是墨菲和迈尔斯(Murphy & Myors,2004)的统计检验力分析。多数情况下通过估计检验力,而不是"猜测"效应大小来计算 N,他们也提供了一个可以让读者通过 d 值或方差百分比来估计 N 的附录。这个附录十分方便,因为元分析家利普西和威尔逊(Lipsey & Wilson,2000,附录 B)提供了一个从各种各样的统计数据中计算 d(或者 d 的等值)的公式,你可以从已发表的文献中找到这个公式。最后我们要提醒你的是,d 所对应的效应大小存在一些惯例:0.20 代表效应小,0.50 代表效应中等,0.80 代表效应大(Lipsey & Wilson,2000[①];Murphy & Myors,2004)。但这并不适用于不常用或很复杂的统计中(如分层线性模型),这时可以查询关于分析的统计文献并请教专家。

设置及器具

描述研究的一般情况,如果是常见的特征就不需要详细描述。在一个实验室实验中,我们需要说明实验室的总体尺寸及每个实验室的装饰,并指出主试与被试在实验室中的位置(如:实验室 9 英尺 × 12 英尺,被试与主试面对面坐在 4 英尺宽的桌子两端)。然而,对于从大学或小学班级里按组收集数据的研究来讲,这些说明是没必要的。另外,由于杂志篇幅比论文的要小,所以准备发表论文时要省略这些细节。如果实验使用了某种特定的仪器,那么就需要详细描述并提供这种仪器的结构图(在正文或附录中)。如果是商品仪器,说明其品牌及型号就可以了。

自变量

我们一般在设计、程序或自变量部分说明如何操作自变量。把这一部分放在哪儿取决于自变量的复杂性:对于复杂的自变量,如干预项目或呈现一组特定刺激,需要一个单独的部分来描述自变量,或者在附录中详细描述自变量(如治疗手册)。

无论把自变量放在哪儿,一些细节都是至关重要的。除了说明如何操作自变量,还需要保证这些变量的整体性。也就是说,阐明实验如何确定对这些变量进行了一致性操作(参见 Billingsley, White & Munson,1980;Kazdin,2003)。这可通过评判员的操作检查(如,对治疗者或实验者的行为进行评价进而保证他们操作的一致性)及实验者的预防措施(如通过单向玻璃来观察实验过程)来实现。如果实验需要指导性的操作(例如,让被试以为自己在一个两人组对抗的游戏中获胜),则要说明用于评估操作是否有效的方法。

有时操作检查很复杂。例如,我们的一个学生想知道如果四年级的女生频繁地

---

① 本书中文版《实用数据再分析法》已由重庆大学出版社引进出版。——编者注

表现出某些对同龄人来说很重要的行为,那么这些同龄人是否会认为这些四年级女生的社交能力更强(Hoier,1984;已发表的文章见 Hoier & Cone,1987)。这些社会行为的增加是自变量,能力的评定等级是因变量。为了保证社会行为确实增加了,赫耶尔对实验过程录了像,由经过训练的观察者对录像进行评分。我们需要在方法部分说明这个评分系统的可靠性,这是很重要的;把操作检查放在自变量部分还是测量部分,取决于对读者来说放在哪儿是最清楚易懂的。

一定要说明谁是主试。大量的文献告诉我们主试或数据收集者不应该知道实验假设,所以主试不应该是自己,除非你采用了某种让自己不知道被试都在哪种条件下的方法。我们一个学生就很聪明地做到了这一点:她想评估一份新问卷的辨别效度,这个问卷用于评估与人格障碍相关的信念。她采用标准的结构性访谈来把被试分组,这个访谈需要花一个小时来掌握并且需要大量的训练。由于经济条件所限,她不想雇人来完成这些访谈,如果她自己作为访谈者则需要保证自己不能知道被试可能有的症状。于是,她让临床医生给她推荐潜在被试,但不告诉她被试可能存在的症状,这样她就能自己进行访谈了。

如果不能避免由于知道被试信息而带来的影响,并且雇不起其他人当主试的话,那么你可以考虑下面的建议。

如果你在大学里,可以请本科生来担当实验主试,并用学分作为酬劳。或者,你也可以和一个与你经济水平相仿的同事交换:你同事作为你的实验主试,而作为交换你去当他/她的实验主试(注意,不可以为了得到社会支持而与这个人详细讨论你的项目!)。最后一个也是最不推荐用的方法是自己当主试,采用能够避免因条件不同而对被试区别对待的程序,这包括让所有被试完成一个关于主试的问卷,或者对实验过程进行录像,并让不知道实验假设的人对其进行评估。

如果实验的实施需要特殊技能或精细的步骤,那么就需要训练主试并保证他们在研究之前及研究过程中能够胜任。这一步骤对于包含干预、动物手术以及使用实验助手(在实验中表现出某种行为)的研究特别重要。

最后,说明一下为了防止混淆而进行的控制程序。例如,给被试呈现一系列录像刺激,说明你是如何控制可能存在的顺序效应的。如果不同的主试实施自变量的水平不同,如何确定影响实验结果是自变量的不同水平而不是主试间的差别?提供这些信息能够让读者相信你的条件只在自变量上存在差别而非其他。当然为了提供这些信息,需要事先找出这些潜在的混淆变量,并消除或评估它们对实验的影响。

因变量(测量)

测量或因变量部分讲的是一些收集数据的工具。第 8 章会讲述相应工具的功能。一旦选择了这些工具,就需要对此进行说明。

如果已经发表了,要说明这些现成工具的名称、作者和发表日期。如果没有发表,在附录中提供此工具的复本。如果这个工具有版权,一定要获得使用的书面许可。除了要说明这个工具在文献中出现在哪儿,还要说明它的形式:给被试解释或复述指导语、描述典型的测验刺激或项目,及其反应选择项(例如,对—错或 1 ~4 个等级)。

虽然你需要将整个工具放在附录中,但仍可以把典型的项目放在一个表格内,这样读者看起来就会很方便。当然,如果使用的是一个非常著名的工具,如明尼苏达多相个性调查表—2,那么就不必详细说明了,因为多数读者都很熟悉它。另外,为了保护测验工具,一些测验中的项目(例如,韦克斯勒成人智力量表—修订版)不

能对外界公布。最后,说明测验是如何计分的,如果有多个分数,指出你想用的分数是哪个。

除了要描述测验的形式特征,还要说明其心理测量属性。在这里,如果只说明信度和效度是不够的,因为它们是相对的。此外,还需要说明你所选择的测量工具可以解决项目中那些你所感兴趣的特定问题。因为某一分数在其他人的研究中是可靠的,但在你的研究中却不一定。

可靠性这个术语太笼统了以至于不能提供太多信息。可靠性或概括性(Cone,1977;Cronbach,Gleser,Nanda & Rajaratnam,1972;Foster & Cone,1995)有各种不同类型。例如,如果要在干预前后评估被试,测量工具的时间稳定性(如再测信度)是个重要特征,而不必太关心它们的内部一致性。此外,如果在前后测验中各组的平均数变化不一致,这时需要关注平均分数的稳定性而不只是测量工具的时间稳定性。类似的还有,两周后的再测能够提供一些关于测量工具的时间稳定性的信息,如果干预需要更长时间,那么就需要查阅文献,看看测量工具在再测时间多于两周时的表现。

除了考查不同类型的信度,还需要说明测量工具的效度。例如,采用自我报告测量时,一定要报告区分效度(Campbell,1960),也需要报告聚合效度。其他类型的效度(如校标关联效度、构想效度、治疗效度及区分效度)的重要性因研究的不同而不同。我们将在第 8 章详细说明。这里要谨记:把重点放在你想要使用的分数上。许多测量工具能够以不同的方式计分,这些分数的信度和效度是不同的。对于不关心的分数可以不对它们的心理测量属性进行讨论。

上述这些建议都与因变量的测量有关。那么自变量呢?如前所述,你应该设法让它发生在预期的水平上。可以在测量部分(自变量部分或程序部分)说明具体操作方法。同样,可以在测量部分或被试部分描述被试。至于这些信息应该放在哪儿,取决于放在哪儿最方便于读者阅读。

此外,如果打算采用人为观察评判计分,需要说明他们是谁(例如标明实验的观察员是本科生)、招募标准、选择及训练方式,以及如何判断他们已达到标准(见 Tryon,1998,直接观察的实践指导方针,多数方法可用于其他需要人类判断的研究中)。另外,说明观察者或评判员间的一致性是多少(我们的最低标准为观察者或评判员至少在每种条件或每个被试的 20% 的数据中做出相同评定)以及是如何计算这种一致性的。科恩(Cone,1999a)和哈特曼(Hartmann,1982)总结了大量用于计算观察者间一致性的不同方法,并提供了一些公式(也见 House,House & Campbell,1981,讨论了其他公式)。最后,说明用于评估评判员或观察者间一致性的程序。最好不要让评判员或观察者知道谁会检查一致性以及检查的时间,因为这会增大一致性估计(Romanczyk,Kent,Diament & O'Leary,1973)。

采用观察或判断等级时,最后需要对分类系统(包括定义)及其程序(包括数据收集表格)进行完整描述。通常情况下,我们在方法部分总结这些,并在附录中提供手册或指导方法及计分表格的复件。如果在研究中修订了测量工具,就不需要前人分数的心理测量属性。在这种情况下,我们建议你进行初步研究。如果在提案会议之前没能收集到初步数据,一定要说明在研究过程中你打算如何得到可靠有效的信息。在通常情况下,这项工作在收集主要研究数据之前就应该做完。理由很简单,若在收集主要研究数据的同时收集信效度数据,如果信效度很差,那么如何解释你的主要研究数据?如果数据不符合预期,这很可能是因为测量工具有问题,而不

是假设的内在逻辑的问题。换句话说,测量数据不能够可靠或者有效地回答研究问题。如果你想对这个问题有进一步的认识,请查阅科恩(Cone,1992)的文章,它对这个问题进行了详细的讨论。

程序

本部分将描述收集数据的步骤(Yates,1982)。如果研究助理根据志愿者名单通过电话或网络通知被试参加实验,那么你需要对此说明一下,并准确描述研究助理在电话中对被试说了什么。如果研究需要邮寄问卷或者其他信息,也要说明其过程。而后,描述被试到达数据收集地点后的步骤,每一步都由谁负责。例如,如果把被试随机分配到每个实验组或控制组,要说明随机化是怎样做的。特别要注意被试对实验条件的反作用及其他形式的被试偏向,如霍桑效应(参考 Ray,2006)。只告诉被试有关参加实验的信息,尽量采用双盲的方法,把实验者偏向降到最低。

为了保证被试的知情同意,需要向他们描述实验程序,包括人类被试的观察过程,并把知情同意书的样式放在附录里。此外,还应描述在实验完成后告知被试相关信息的过程,并形成书面文字放在附录里。这里要提醒你的是,不要忘了被试有权了解研究的结果,你可以在知情同意书里设置一个选项,如果他们想了解研究结果可以在知情同意书上注明(同时也可以让被试留下他们的联系方式)。

完整细致地叙述呈现给每个被试的指导语是什么,如果各组的指导语不同,还应说明其不同之处,并把所有指导语的复件放在附录里。

通过录音向被试呈现指导语能够保证指导语的标准化,并且有助于避免实验者的体力不支、交流倦怠、不耐烦,或者其他特殊反应。可以回顾之前我们在整体检查部分的建议。

最后,如果研究设计需要治疗后评估或事后评估,一定要描述其评估过程:评估是在什么时候进行的?需要召回被试吗?是通过电话还是邮件?

熟悉心理学中的研究道德

这一部分会详细说明研究道德。希望读者通过对这一部分的阅读,能够对怎样以一种符合道德的方式进行研究有进一步的认识。我们首先简要说明心理学研究的道德原则,然后会提供一些实施这些原则的建议。

心理学家的专业行为——无论是临床心理医生还是研究者,无论是初出茅庐的新手还是有经验的老手都要遵守由 APA(2002)发布的道德原则和标准。附录里有标准里的2.05,3.04,3.08,4.01,4.02,4.04,4.07,5.01,6.01—6.02,8.01—8.15,这些都与研究有最直接的关系,它们的本质在于,当心理学家进行研究时,他们能够按照这些标准做,并把被试的利益放在最重要的位置。

在履行尊重被试利益的责任时,道德守则及机构审查委员会的要求给我们如下建议,这些不是具体的原则和标准,只是我们的经验之谈。

1. 评估研究的道德可行性。
2. 评估被试面临的危险程度。
3. 保证研究中你和其他实验者的行为符合道德。

4. 被试参加时，要签订一个清晰、公平、知情并且自愿的协议。
5. 除非万不得已且公平的情况下，不得欺骗或隐瞒。
6. 被试有权在任何时候退出实验。
7. 避免对被试造成可能与研究程序有关的任何身体伤害，危险或不舒适。
8. 避免对被试造成可能与研究程序有关的任何情绪伤害，危险或不舒适。
9. 完成数据收集后要对被试报告结果。
10. 纠正因为参加研究而给被试造成的任何不良后果。
11. 对在研究中收集的有关被试的信息进行严格保密。

如果研究采用的被试不是人类，应恪守道德原则 8.09，“人文关怀及研究中的动物使用。”在这类研究中需要注意下面几点。

1. 符合有关获得、照料及处理动物的法律和专业标准。
2. 确保研究者经过专门的训练：能够照顾动物，能够监督动物的使用，并最终对它们的健康、舒适及人文关怀负责。
3. 确保所有的研究者都了解研究的方法，并对如何照顾、保护及处理特定动物经过专门的训练。
4. 把发生在动物身上的任何不舒适、感染、疼痛或疾病降到最低。
5. 尽量不要采用会产生压力、疼痛或困苦的程序，研究目标是通过实践价值、教育价值及科学价值来体现的。
6. 在适当麻醉后再进行手术，在手术中和手术后要采用能够避免感染和把疼痛最小化的程序。
7. 终止动物的生命时，要采用快速适当的程序，并把疼痛降到最小。

在研究中实施道德标准

如果研究者的行为符合标准，那么此研究就是道德的、负责任的。本部分将提供一些相关建议。如果你想了解更多的细节，请查阅《人类被试研究中的道德》(*Ethics in Research With Human Participants*)(Sales & Folkman,2000)，或是在纳吉(Nagy,2005)的《通俗英语中的道德：心理学家的个案记录簿》(*Ethics in Plain English:An Illustrative Casebook for Psychologists*)一书的第 8 章中也有简要的介绍。

道德可接受性

研究的第一项要求就是保证在道德上是可接受的。达到这项要求的一个简单的方法就是遵守我们提议的那些原则。遵守上述原则后研究还能进行吗？如果答案是肯定的，那么在道德上就是可接受的。如果答案是否定的，在道德上仍然可能是可接受的，只是需要采取其他措施来确定研究在道德上是可接受的，并向自己和他人保证它在道德上是可接受的。

道德可接受性要特别关注研究中的欺骗和隐瞒，这是因为欺骗和隐瞒只有在“完全需要和合理”的情况下才可以使用。举个例子，设想一个需要被试与他人竞争的研究，为了精确控制其他个体的行为，实验者可能需要编一个实验程序来与真正的被试竞争，而告诉被试的是与他/她竞争的人在另一个实验室中，他们能够交流

但却看不见彼此。被试不知道在其他房间里的人其实是计算机。

很明显,在这个实验中存在欺骗。这不符合禁止欺骗的规定吗?你能够说明这个实验的重要性并且如果不欺骗被试就不能完成吗?你能否采用一点儿有创造力的手段,或者是先进的统计方法来避免频繁的欺骗(Kazdin,2003)?例如,在上面的例子中,可能告诉被试他们在与计算机竞争与告诉他们在与真人竞争所得的结果会差不多。如果研究重点是关注竞争,且计算机竞争者与真人竞争者是等价的,那么就没有必要欺骗了,干脆告诉他们在与计算机竞争好了。

不幸的是,就像上面这个例子所说的,如果需要做大量的工作来保证有争议的研究方法的道德性,这不仅会耗费很多时间并且常常会超出新手的能力范围。如果想通过完成项目来获得学位,那么即使不进行试验研究也没有关系。

道德的可接受性促使你在当前可行的研究领域中进行研究。绝大多数的研究都是遵循 APA 的道德原则的,不符合这一原则的少之又少。可以参考文献中已有的先例来向你自己、机构审查委员会成员、论文答辩委员会成员及其他人保证你的研究符合道德原则。当然,只是参考已有的文献是不够的,需要陈述每一条道德原则,并向论文答辩委员会和学院的机构审查委员会保证你正在向道德标准努力。

设计知情同意程序

在符合道德的研究中,被试的知情同意是一个很重要的要素。知情同意是一个过程,包括两部分:告诉他们在研究中都需要做什么并获得他们的书面同意。

知情同意的一个环节是给被试一份知情同意书。如果被试不到 18 岁,需要把知情同意书给他/她的父母或监护人。同意书里的具体内容因地点而异,但要符合机构审查委员会的制度变动。根据我们对一些机构审查委员会以及美国心理学协会道德准则的理解,下面是任何同意书中都应具备的基本元素。

1. 描述研究及其目的;
2. 向被试提供实验需要的信息;
3. 描述被试需要做的事及需要多长时间;
4. 描述对被试来说存在的潜在危险和好处;
5. 说明被试是自愿的并且在任何时候都可以不受惩罚的退出;
6. 告诉被试所有的数据都会得到保密,并描述不能保密的所有情况;
7. 向被试提供研究责任人的联系方式,告知被试可以向他们索取更多有关研究的信息;
8. 向被试提供投诉负责人的联系方式,如果被试对于参加这项研究感到不满,可以与他们取得联系;
9. 说明被试是否会获得酬劳及得到什么酬劳;
10. 如果研究考察的是一种实验处理,描述研究中的其他选择;
11. 向被试说明他们有权获得研究的结果;
12. 留出被试及研究者用于签名的空间。

专栏 7.1—专栏 7.4 提供了在不同情境下采用人类被试的同意书样本。专栏 7.1 是一个成人同意书——约瑟夫·塞韦里诺的博士论文采用模拟方法考察了厌恶犯罪效应。让男同性恋及陌生男人听录音带并“出声思维”,录音带描述的是有

或者没有反同性恋语言的人身攻击。实验中,塞韦里诺特别关注那些有严重创伤历史的男性。因为,实验可能会再次激活他们的创伤,所以他在招募被试时排除了那些有特殊痛苦创伤历史的个体。另外,他还告诉被试这个录音带包含图形语言。

专栏 7.2 描述的是父母如何同意他或她的孩子参加研究。由于这项研究是在学校中进行的,所以需要把同意书送到家里让父母签字。专栏 7.3 是在同一项研究中让未成年被试签的同意书。专栏 7.4 是另一项研究中用于更小儿童(4—6 年级)的同意书。这份同意书里采用的是适合儿童的语言。

在与儿童或其他在法律上不能够签同意书的个体打交道时,他们的父母或法定监护人一定要代替他们签同意书。在简单描述你的研究之后,这些潜在的被试应该有机会同意或拒绝参加研究。另外,也应该要求儿童签同意书(见专栏 7.4)。

专栏 7.1

成人参加研究的同意书样本

阿连特国际大学,圣地亚哥分校
10455 Pomerado 路
圣地亚哥,加利福尼亚州 92131
机构审查委员会,(555)555-5555

男性对犯罪的反应

欢迎参与这项研究。在您同意参加实验之前,我们希望您能阅读下面的内容,可以随时提问,进而确保您能明白在研究中都需要做些什么。

研究者

实施这项研究的学生是约瑟夫·塞韦里诺,文学硕士(MA)。监督这项研究的教师是沙伦·福斯特,博士。

研究目的

本项研究的目的是为了更好地理解人们对各类犯罪的反应。在经历犯罪被害时,个体试图处理他们的经历,这会产生大量的思维过程。本研究的目的就是详细描述经历不同类型伤害的个体的思维过程。

参加研究的时间及被试人数

如果您同意参加实验,这大概会花费您 60 ~ 90 分钟的时间,并且如果你不想告诉研究者你在研究中的经历可以随时退出实验。研究大概需要 80 人。

研究程序

如果你选择参加这项研究,首先需要填写两份问卷,问卷涉及你在过去可能有的创伤经历。对于那些在过去有创伤经历的人来说,填写这样的问卷可能会勾起他们伤心或烦恼的记忆。一些参加者在这项研究中可能只需要填写这些问卷。除了这两份问卷,还会要求您听一段有关犯罪的录音,并想象自己是受害者。进而可能会让你描述如果这个犯罪事件发生在您身上您会怎么想。我们会对您的描述进行录音。您能看到所有的记录工具,并且如果您不同意记录那么我们就不会记录。除了言语反馈,您还需要填写两份或多份有关于自身反应的问卷。

(待续)

专栏 7.1 (续)

风险

参加本项研究的风险在于您将会接触到存在潜在烦恼、唤起和/或焦虑的材料。研究内容包括具有贬义的语言和有关暴力行为的描述,这可能会冒犯、侮辱、惊吓和/或干扰到您。但研究中的暴力行为不包括性行为或严重伤害。您可以在任何时候要求停止录音播放并可以无条件地从研究中退出。另外,您需要回忆过去的创伤事件,这可能会勾起让您伤心的记忆。

研究中的好处

您也许不能从这项研究中得到任何直接的好处,但研究者希望您能从研究中多学些可能在以后会有助于他人的东西。

研究中的其他选择

如果选择参加这项研究,那么请遵守上述要求。您可以选择不参加,并且也可以在任何时候无条件地退出这项研究。

保密

我们将尊重您的隐私权,任何有关您身份的信息都是匿名和保密的。您对所有问卷的回答、录音及书面测量都是用数字进行编码的,并且只有主要研究者可以看到真正的名字。在之后任何材料上都不会出现有关身份的信息。在没有经过您或法律规定的允许时我们不会泄露在本研究中收集到的有关您身份的信息。在没有泄露您身份的前提下,本研究结果可能会发表在专业杂志上或者在心理学会议上报告。在某些情况下,数据可能会被法院传唤。

与研究有关的疑问

如果您有什么疑问可以随时联系约瑟夫·塞韦里诺,电话是(555)555-5555。此外,在签这份同意书之前您可以提出任何问题。

被试获得的酬劳

实验完成后可获得一张免费电影票。另外,您还有机会参与抽奖,奖品是 Game Boy 或便携式电视。

先前参加过的研究

在过去的 3 个月中我参加过以下研究:

____

____

参加权及退出研究

自愿参加本研究,可以拒绝参加或者在中途退出。你仍然会得到酬劳,并且不会受到任何惩罚。我们会向您解释研究中的所有重要细节。如果还有问题,研究者会很乐意为您解答。

签名及确认

若在下方签名,则意味着我已经阅读了上述信息并且可以问有助于参加研究的问题。在改变主意之前我同意参加研究。我已收到这份同意书的复本和被试权利法案的复本。我并没有因为签署这份同意书而放弃任何权利。

被试签名______________________ 日期______________________

见证人签名______________________ 日期______________________

调查员签名______________________ 日期______________________

注:转自 Gay and Straight Men's Attributions to Simulated Antigay and Nonbias Victimization, by Joseph Severino. Unpublished dissertation proposal, 2002, California School of Professional Psychology, Alliant International University, San Diego. 使用经过作者授权。

儿童同意书类似于成人同意书，但要用简单的语言书写。我们认为，如果法律监护人或潜在被试拒绝签同意书，那么就应该自动把此被试从研究中排除。

除儿童外还有其他群体需要特别关注，以保证他们是自愿参加的。包括精神迟滞的个体、囚犯、无家可归者、现役军事人员、雇员及学生。另外，可能一些人由于能力有限，不能够理解自己同意的是什么，如学生。而对于另一些人，需要预防间接的胁迫。例如，雇员们会因为主管的要求而违背自己意愿，假装自愿参加实验。

专栏 7.2

父母同意孩子参加研究的同意书样本

亲爱的父母或监护人：

您是否同意您的女儿参加一项研究？这项研究是关于您的女儿如何影响他人在饮食、体重和体形上的态度和行为。研究题目是“同龄人关注并影响与身体形象有关的问题”。这是博士学位论文的一部分。具体来说，我想研究身体形象及节食对于女孩儿及她们的朋友的重要程度，及其与她们饮食习惯的相关程度；她们如何相互比较，及她们对自身形象的感觉。

研究内容如下：来自圣地亚哥高中的 150 名女生会填写几份问卷。在这期间，会给每个女生一份她的班级中参加研究的同学名单，并指出自己常和谁一起玩及自己对她们的了解程度。女生们需要填写几份有关与朋友的关系及朋友的态度和行为的问卷。她们也需要报告对自身形象和饮食行为的感觉。另外，我们会私下测量每个女生的身高和体重。时间为 60 ~ 70 分钟。

可能存在的好处和风险：我会尽量不让您的女儿错过重要的课程。不过，她可能需要根据老师的日程来补一些功课。您的女儿不会因为参加研究而产生严重的不适。尽管她可能会由于回答一些问题而有些许的不舒服——她需要讨论自己的饮食习惯及对自己体重和体形的感觉，但是如果她不想回答那么就可以不回答。参加实验的好处是可以更好地理解同龄人如何互相影响彼此的饮食习惯和身体形象，这有助于预防和治疗饮食问题。我们会为所有被试提供偶尔会影响她们饮食问题的信息。作为答谢，每个女孩儿都会得到一张电影票。

本实验自愿参加。如果您不希望她参加，她不会因此受到任何惩罚，并且她可以随时中途退出或者拒绝回答任何问题。本研究得到了阿连特国际大学保护人类被试机构审查委员会的同意。

保密。我们会对所有信息保密。只有研究者可以看到问卷。所有问卷都是匿名的。名字用识别号码代替。

我们会告诉每个女生她们的数据是保密的。如果研究中的某个被试正在做对自己或他人有潜在危险的事情，那么我会调查这件事并且有可能报告给有关人员，我会让每个被试都明白这一点。当然，发生这种情况的可能性很小，但是有必要在您做决定之前告诉您保密的有限性。

有问题吗？无论您是否同意您的女儿参加，我都非常感激您的回复，这样我就会知道这份同意书已经到了您的手里。您可以留下附件作为记录。如果有任何问题，可以随时给我打电话，埃琳·安德森 (555)555-5555。如果您愿意的话我可以事先让您看一下问卷。本研究的监督员是沙伦·福斯特博士，电话是(555)555-5555。阿连特国际大学的机构审查委员会也能够回答有关被试在研究中的权利的问题。他们的电话是(555)555-5555。另外，如果您选择参加研究，那么您和您的女儿会收到被试权利法案的复本。权利法案是一个小册子，描述了您女儿作为被试的权利。值得注意的是，小册子中的一些规定并不适合本研究。

（待续）

专栏 7.2 （续）

感谢您的阅读。

诚挚的，
埃琳 K. 安德森，文学硕士
博士生
加州大学心理学专业

选择下面的合适项并让您的女儿把这份同意书带到学校。

□我已经阅读并理解了这封许可信。如果我女儿愿意的话我同意她参加这项研究。

□我已经收到安德森女士让我登记的来信的复件。

□在我决定我女儿是否参加之前我需要更多的信息。请给我打电话__________。

□我不希望我女儿参加。

父母的签名/日期____

孩子的签名____

请让你的女儿把这份同意书带到学校。谢谢!!

注:转自 Peer Influences on Body Image Related Concerns:The Role of Social Comparison, by Erin K. Ferma. Unpublished dissertation, 2005, California School of Professional Psychology, Alliant International University, San Diego. 使用经过作者授权。

专栏 7.3

青少年同意书样本

研究题目:同龄人影响与身体形象有关的问题

调查员:埃琳·安德森,文学硕士(555)555-5555

沙伦·福斯特,博士(555)555-5555

我自愿帮助安德森女士做这项研究。这项研究的目的是为了更好地理解身体形象和饮食习惯对女孩儿及她们朋友的重要性。

如果我决定参加,在一两个学期内这会花费我 60～70 分钟的时间。在这期间我需要在班级女生名单中指出谁是我最好的朋友。另外,我需要填写几份问卷,这些问卷与我和朋友之间的相互影响、我朋友的态度和行为及我对自己形象和饮食行为的感觉有关。也会测量我的身高和体重。我知道所有的信息都是保密的。

我知道我的所有信息和回答都会得到保密。我知道如果我因为参加研究而错过了某一节课的一部分,我需要把错过的功课补上,并且如果研究在放学后进行,那么我可能会错过一部分课外活动。我也知道通过参加这项研究,我可能更好地理解我是如何影响我的朋友的,以及这如何影响她们和我的态度和行为。另外,我会得到一张电影票作为答谢。

研究人员已经向我解释了这项研究并且我可以问任何有关这项研究的问题。我知道如果我不想参加的话就可以拒绝填写问卷,且没有人会因此责怪我。我可以中途退出,且可以跳过我不想回答的问卷。我知道我的所有信息和回答都会得到保密。然而,我也知道如果我正在做有潜在危险并会伤害到自己的事情,那么研究员就会公开信息并且告诉我的父母。我已经阅读了这份同意书,明白了这项研究,并且同意参加。

（待续）

专栏7.3 (续)

学生签名/日期____

研究员/日期____

注:转自 Peer Influences on Body Image Related Concerns: The Role of Social Comparison, by Erin K. Ferma. Unpulished dissertation, 2005, California School of Professional Psychology, Alliant International University, San Diego. 使用经过作者授权。

专栏7.4

孩子参加研究的同意书样本

我明白在这项研究中我需要观察女孩儿们对其他女孩儿的想法、感觉和行为。我会在一个小组中与其他女孩儿及采访者会面,这个会面大概需要60分钟。在会面期间,我会回答一些问题,这些问题与我所熟悉的女孩儿的行为有关。我也会读一些有关女孩儿们在一起做的事情的故事并回答与这些故事有关的问题。

我知道如果不想回答的话我可以不回答任何问题,并且可以随时回到我的班级。我可以在任何时候退出。这不会影响我的成绩,并且没有人会因此责怪我。如果在研究中问一些让我感觉不好的事情时我可以告诉莱弗蒂女士。如果我决定参加这项研究,那么在会面后我会得到一个小礼物。

我知道我的回答是保密的并且采访者不会向任何人透露我说过的话。我也会对自己的回答保密并且不会告诉其他小朋友。

如果我有什么问题,可以现在问或者让我的父母或老师给莱弗蒂女士打电话(555)555-5555或给福斯特博士打电话(555)555-5555,分机是555。

如果我已经阅读且理解了这份同意书并且同意参加,那么就在下面签名。

儿童的签名_______日期____

采访者的签名_______日期______________________________________

注:转自 The Relationship Among Relationally Aggressive Behavior, Emotion and Social Cognitions in Preadolescent Females, by Cambra Laverty Finch. Dissertation, 2001, California School of Professional Psychology, Alliant International University, San Diego. 使用经过作者授权。

要求大学生参加研究以便达到某门课程的标准是一种受到限制的胁迫(Neuman, 1997),特别是对那些不想参加可是又没有其他选择的学生。可以从特定群体中寻找援助,如从你的主管或学校的机构审查委员会中。

另外,如果研究中涉及少数民族群体那么就需要考虑一些特殊的民族问题。

如果不是有目的地研究特定群体,那么你可能很希望自己研究中的被试来自于各种种族和民族群体。你需要重点关注很多事情,像是如何进行分组,如何把你的想法和方法应用到各个群体中。当研究样本中只有一小部分是少数群体时,需要避免不恰当的概括,此时不可能检验出结果是否适用于不同群体(见 Foster & Martinez, 1995)。当研究主要关注少数民族群体时,需要考虑一些额外的特殊问题。费希尔等(Fisher et al., 2002)广泛探讨了研究儿童时如何生成并使用具有文化气息的

且尊重人的同意书和保密程序。这个也适用于成人研究。

注意同意书的阅读水平。确保这个阅读水平不超过被试最低的阅读熟练度。研究表明,以往许多同意书都大大超过了被试的阅读水平。例如,一些医学研究同意书,要求的阅读熟练度要比病人的实际平均水平高出 5 ~ 6 级(Hochhauser,1999)。计算同意书的阅读水平很容易,许多文字处理程序都包含阅读水平计算器。例如,Word™,即写这本书所采用的程序,对前两句话的 Flesch-Kincaid 等级水平的估计是 8.7。

被试首次参加实验时,签署同意书常常是第一道程序。如果选用的是在校学生,需要把同意书寄到他们家里,家长签好字后再让学生带回来。无论家长是否同意都需要回执。这能确保家长至少收到了同意书。

一些研究者(如 Clark,Lewinsohn,Hops & Seeley,1992;Roberts,Lewinsohn & Seeley,1991)成功使用了被动同意程序。在克拉克等人(Clarke et al.,1992)的研究中,如果家长不同意孩子参加,那么他们需要把拒绝卡片返回给研究者。当然,在孩子们了解程序之后他们也可以拒绝参加。这种程序比主动方法的参加率高。如果采用被动策略,首先需要得到保护人类被试的机构审查委员会的同意,并解决与你的负责人存在的潜在分歧。需要注意的是被动同意程序有可能存在争议。一个研究者在一个采用有争议程序的研究中使用了这种方法来研究学龄儿童,收集完数据后家长提出了抗议,最终研究者不得不毁掉上百个被试的数据。被动同意程序也存在教育和种族偏见,即研究者会更多地报告受过较少教育的个体和少数民族(Dent,Galaif & Susman,1993)。同意书里是否包含其他的项目取决于你所在学校机构审查委员会的偏好。

考查被试的风险

考虑被试的潜在危险是道德研究及获得机构审查委员会同意的关键部分。必须评估所有的危险,无论是身体上,情绪上,还是其他方面。如果采用和前人类似的程序,那么可以从已有文献中寻找答案。极少数的研究会直接考察研究程序的危险性(参见 Bell-Dolan,Foster & Sikora,1989)。在设计自愿同意书时可以考虑这些研究结果。多数情况下研究者们没有直接考查潜在危险,而是从文献中收集信息。这需要阅读大量文献,并指出是否有身体、情绪或其他类型伤害的报告,以"许多采用这些程序的研究中"的形式或其他能够表现同样效果的文字来进行。

根据我们的经验,新手有时会低估危险,但更常见的问题是他们高估被试因为参加研究而苦恼的可能性。例如,我们看过机构审查委员会决议草案中关注被试因为回答有关他们的坏习惯、负向情绪之类的问卷而感到的不适。需要记住的是"最小风险"的标准是被试在日常身体或心理考试中经历的危险程度。人们在和心理学家谈话时常常会被问及个人信息,所以我们认为报告正常的日常信息不会导致过度的不安。

总之,如果问你多长时间会产生关于自己的消极思想,你会因此而感到严重不安吗?被试真的有可能感到不安吗?同时,涉及高度敏感的信息(如关于一个人的性习惯,使用非法药物等)可能会产生更大的问题或导致更大的危险。如果有疑

问，可以请教专业人士，收集他们对潜在危险的看法。

了解机构审查委员会的操作

IRB 审查研究的步骤通常是会公开发表的，你应该借复本来读一读。他们会提供在收集数据前为了使研究通过他们检查所要遵守的步骤。

由于 IRB 通常也要检查预实验，所以在你设计实验时就要注意 IRB 会议的日期及提交申请的最后期限和原则。如果会议每月开一次，错过一次会议意味着你的研究要推迟一个月才能开始。弄清楚你的研究属于哪一类，具体要求是什么。或许你的研究计划可以免除 IRB 的审查或者可能进行外审。当数据来源于档案或你正在采用相对良性的调查，或者观察大众行为，或者评估正常的教育实践，那么就不需要 IRB 全面审查。请教你的导师或论文答辩委员会成员。他们在当地 IRB 程序及政策上很有经验，并且一定能够在操作上给你一些明智的建议。

知道 IRB 的批准时间比知道提案会议的时间更重要。一些学校要求学生在提案会议之前获得批准。这是因为研究偶尔会有重要的改变需要向 IRB 报告，这可能会影响到研究的基础部分，包括基本问题和假设。答辩委员会成员们不想再开会来考虑这些改变，其他学校希望提案会议会对原有方案进行重大改变以至于 IRB 将会评估一个与原有方案不同的提案。这些学校 IRB 的批准很有可能在提案会议之后。

无论是在提案会议之前还是在之后获得通过，如果研究存在一点儿争议，那么都要事先预料到这很难通过 IRB 程序。与此相关的一个有趣的例子来自于沙伦·福斯特几年前有关儿童社会技能领域的研究。程序要求儿童点名指出 3 个自己最喜欢和 3 个最不喜欢的同龄人。IRB 的一位成员坚持认为这存在潜在危害，提名会吸引被试对不喜欢儿童的注意并且会更消极地对待他们。福斯特说虽然只有一项研究直接探讨了这个问题，但没有数据表明这个程序有危害。由于数据有限，福斯特决定招募这位 IRB 成员的助理来设计一个实验以检验提名程序是否对儿童产生了负面影响。特别关注这个问题的 IRB 成员同意了这个测验，一个学生实施了这项研究，并以此作为自己的硕士论文。顺便提一句，提名程序没有产生负面影响（Bell-Dolan et al.，1989），IRB 成员们后来也没有反对过这个程序。

持续监测研究道德

设计一个在道德上可接受的研究方案并让他人（如 IRB 和论文答辩委员会）信服是准备收集数据的基础。然而，确保研究的道德行为并不是到这儿就终止了，事实上这只是开始。开始收集数据时，需要确保自己及任何帮助你实施研究的人一直以符合道德的方式进行研究。现在应该考虑如何保密，如何确保被试完全了解这项研究，等等。考虑一下自己将如何操作及如何把它放进提案中。我们在第 10 章讨论实施研究时会再谈论这个话题。

关于研究道德的额外细节，请参阅《APA 写作手册》及心理学家有关研究的道德原则（见本卷附录）。如果研究不属于心理学领域，那么也应该遵守类似的道德原则。本章探讨的基本方法原则和道德原则超过了学科界限，理解它们与制订方法一样重要。

补充资源

研究设计

Barlow, D. H., & Hersen, M. (1984). *Single-case experimental designs: Strategies for studying behavior change* (2nd ed.). Elmsford, NY: Pergamon Press.

Hayes, S. C., Barlow, D. H., & Nelson-Gray, R. O. (1999). *The scientist practitioner: Research and accountability in the age of managed care* (2nd ed.). Boston: Allyn & Bacon.

Kazdin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Boston: Allyn & Bacon.

Kerlinger, F. N. (2000). *Foundations of behavioral research* (4th ed.). New York: Harcourt Brace.

McBurney, D. H. (2004). *Research methods* (6th ed.). Belmont, CA: Brooks/Cole.

Miller, D. C. (2002). *Handbook of research design and social measurement* (6th ed.). Newbury Park, CA: Sage.

Ray, W. J. (2006). *Methods: Toward a science of behavior and experience* (8th ed.). Pacific Grove, CA: Brooks/Cole.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Solso, R. L., MacLin, M. (2002). *Experimental psychology: A case approach* (7th ed.). Boston: Allyn & Bacon.

Weisberg, H. F., & Krosnick, J. A., & Bowen, B. D. (1996). *An introduction to survey research, polling, and data analysis* (3rd ed.). Newbury Park, CA: Sage.

检验力分析

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

Murphy, K. R., & Myors, B. (2004). *Statistical power analysis* (2nd ed.). Mahwah, NJ: Erlbaum.

研究伦理

Cone, J. D., & Dalenberg, C. (2004). Ethics concepts in outcomes assessment. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (3rd ed.). Mahwah, NJ: Erlbaum.

National Institutes of Health. (2005a). *Guidelines for the conduct of research involving human subjects at the National Institutes of Health*. Retrieved November 7, 2005, from http://ohsr.od.nih.gov/guidelines/graybook.html. [美国国立卫生研究院关于采用人类被试的指导方针的网站]

National Institutes of Health. (2005b). *Regulations and ethical guidelines*. Retrieved November 7, 2005, from http://ohsr.od.nih.gov/guidelines/45cfr46.html. [采用人类被试的实际联邦准则;这个网站提供了采用儿童、囚犯、孕妇及新生儿作为被试的指导方针。]

思考与练习

撰写方法部分

□方法部分的组成

—设计

—被试

—设置及器具

—自变量

—因变量(测量)

—程序

□遵守心理学的研究道德

—阅读道德准则

—撰写同意书

—查找并完成 IRB 提案

—提交 IRB 提案,修改,获得批准

—监控研究过程,确保道德准则的全程执行

8

测　量

本章将讨论如何对自变量和因变量进行操作。其中一些操作会采用测量工具。本章会谈谈好的测量工具所具有的特征及在哪儿能够找到这些测量工具及其相关信息。另外,还会提供一些在没有确定心理测量分数的适当性时的建议。

操纵变量

如果要准备一篇论文或博士论文,应该知道概念的操作定义是什么。如果不知道,那就看一篇有关研究设计的好文章(如 McBurney,1990;Ray,2006;其他建议可看本卷第7章后面的清单)及相关的评论。

在论文的方法部分,要描述自变量和因变量的操作定义。例如,如果要考察处于高、中、低社会经济地位的家庭在凝聚力上的差别,应该如何确定社会经济地位呢?如何测量凝聚力呢?如果要考察糖尿病儿童和患有其他疾病儿童在护理上的差别,应该如何确定被试糖尿病的严重程度?如果预测护理取决于护士的已有经验或有关糖尿病的知识,如何评估这种经验或知识呢?再举一个例子,假设想知道儿童时期受过性骚扰的成人是否比没有受过性骚扰的成人产生更多的心理问题,应该如何定义和评估骚扰?会不会把不同形式的骚扰混为一谈?是否需要区分虐待和骚扰?如果需要,又如何区分呢?如果要研究在某个变量上的种族差别,应该如何定义种族?如何处理混血个体呢?是否考虑融入主流文化的程度?如果答案是肯定的,那么如何评估呢?如果打算把具有高、低"评价焦虑"的个体置于具有高、中、低压力的公共演讲情境中,那么如何操纵压力?同样重要的是,如何确定操纵是否成功?如果想描述在快餐店中肥胖顾客是否比非肥胖顾客吃更多的碳水化合物,那么又该如何进行观察呢?

采用某种客观的工具测量,就可以解决上述大多数问题。这个工具可能很简单,把被试划分为各个社会经济水平或种族群体。也可能很复杂,明确指出了评价焦虑或一些其他人格特征的指标。不管它属于哪种类型,操作指标都需要具备一些能够让科学界接受的品质。

了解潜在工具的重要特性

科学研究中使用的测量工需要具备:①明智选择;②对于特定研究群体及测验分数来说具有某种心理测量特性(即信度和效度);③足够直接。明智选择是指测量适合你所感兴趣的变量。例如,如果关注婚姻关系中的冲突,那么测量应该能够反映出这种冲突。因此,在选择测量工具之前要精确定义变量。

冲突是一般用语并且可以指向多种行为,可能包含与配偶的相互作用。例如,在婚姻关系中一个人可能有很多冲突的想法。如果关注的冲突与配偶间的相互作用有关,那么是言语冲突、身体冲突,还是二者兼而有之?如果明确知道要评估什么,那么就会很容易选择出最合适的测量工具,这对于自变量和因变量都适用。焦虑并不都一样,认识它有许多不同的方式,种族也一样。此外,不要被测量工具的名称所蒙蔽,必须检查它的内容以确定它所评估的事物是否与研究中的定义一致。

也应该明确自己是对行为本身(包括运动动作、思维和感觉经历及生理反应)感兴趣还是对它所表征的假想建构感兴趣,这个区别决定着如何评估测量数据(Foster & Cone,1995)。如果关注被试的行为,那么测量应该反映出被试在真实世

界中的行为。如果关注假想建构,那么测量是否反映事实就不太重要了,重要的是它和其他能够测量假想建构的测量工具适当相关。由于相对于行为评估,多数学生更熟悉假想建构评估,所以我们会花更多时间来讨论行为评估。

如果对行为感兴趣,那么测量工具越直接越好。在行为能够自然发生的时间和地点,用测量工具评估感兴趣的行为(Cone,1978)。在上述婚姻冲突的例子中,研究者可能关注言语冲突。评估言语冲突的方法之一是在访谈中问被试,其他评估方式包括:适当的自我报告,让其他非常了解这对夫妇的人评价他们在言语冲突中的行为,让这对夫妇自我监控(即观察并记录他们在家时的冲突),或者让训练有素的观察员观察这对夫妇并记录所有言语冲突。

可以把这些评估类型排列成连续体,以此来表示所感兴趣的行为与特定评估方法所反映的行为之间在形态、时间及空间上的相似程度。

研究者通常希望被试的反应尽可能的自然,希望能够把数据应用到现实世界中,希望自己的研究能够有尽可能大的外部效度。直接测量能够提高外部效度,专家并不认为通过被试自我报告得到的数据会比直接测量得到的数据质量低。

直接测量需要花费大量时间,让训练有素的观察员在自然环境中进行观察并不容易。直接观察需要花费大量的时间开发并标准化观察代码,寻找并训练观察员,进入到被试的自然环境中,隐蔽地观察他们并把数据整理成方便使用的形式。对比直接观察所需要的时间与开发并实施结构性访谈所需要的时间,即使访谈需要训练访谈者,对访谈进行录音,并把录音整理成书面记录,这仍比直接观察方法所需要的时间少得多。

自我报告测量需要的时间相对更少一些。访谈法需要训练访谈者,转录,并对书面记录进行评分,为什么不让被试从描述他们行为的有限陈述中进行选择呢?为什么不呢,确实,这是许多研究者、新手及其他一些人的观点。事实上,自我报告方法是进行主观测量研究的首选评估形式。在之前的例子中,可能让夫妇们在他们是否"在钱的问题上争吵","在他人面前提高声音分贝"等这样的陈述上回答**是**或**否**。这样做的重要前提假设是夫妻们所报告的与行为的其他重要形式存在某种程度一致性。然而,考察人们言语与实际行为一致性的研究表明这个假设常常是错的(Bellack & Hersen,1977;Shiffman,2000)。此外,回顾性的自我报告易受很多偏向因素影响,并且这些偏向因素多数是正常认知加工的产物(Stone et al.,2000)。

这不是说自我报告方法得到的分数与采用其他测量方法得到的分数不相关。它们常常相关,特别是与其他自我报告方法得到的分数。但是用自我报告代替更直接的评估方式并没有得到太多人的赞同。确实有许多研究者公开承认自我报告可能反映不出某些行为,然而只有很少的研究进一步考察了这一点。

可以考虑其他方法而不仅仅依赖自我报告方法。例如,观察者等级评定的客观性更好一些,这是因为观察者提供的数据更不易受主观影响。然而,这仍然是间接的测量,因为这依赖于其他人对发生在某一时间和地点的行为的报告。另外,在记录时存在一些问题,其中一些在自我报告中也存在,包括不能正确回忆、光环效应、慷慨偏向、对结果的不同解释等(如 Cairns & Green,1979;Cronbach,1990)。

自我观察(有时也称"自我监控"或"生态瞬息评估")是另一个选择。这比访

谈、自我报告及他人的等级评定更直接,这是因为要求被试在所感兴趣的行为发生的同时进行观察和记录。虽然自我观察因为评估者与被评估者是同一人而缺少客观性,但是可以通过让被试只关注明确的反应来减轻主观性。有时可以判断自我观察的准确性,就像判断其他人观察的准确性一样(具体做法可参考 Johnston & Pennypacker,1993)。采用个人数据帮手,如手写随身电脑进行自我观察,已经被广泛应用于日常活动采样、情绪等,特别是在行为医学领域(Smyth & Stone,2003),及青少年研究中(如 Whalen,Jamner,Henker & Delfino,2001)。科恩(Cone,1999b)在 1999 年的《心理学评估》(*Psychological Assessment*)杂志中发布了一系列有关自我观察的文章。Stone,Kessler 和 Haythornthwaite(1991),Foster,Laverty Finch,Gizzo 和 Osantowski(1999),及 Christensen,Barrett,Bliss-Moreau,Lebo 和 Kaschub(2003)都描述了在设计自我观察系统时的决策和问题。

最后,我们建议采用直接观察来评估变量。如前所述,这确实是个高成本的程序。但是有很多方法可以减少直接观察的成本(Cone,1999a;Foster,Bell-Dolan & Burger,1988)。例如,可以记录(录像或录音)行为用于以后分析,而不是让观察者到真实的情境中收集数据。或者观察者可以透过实验室的单向玻璃观察被试,虽然这种方法把行为从自然背景中分开,但它可能仍然比一些不太直接的方法所产生的数据质量更高。录像或录音记录下的行为相对永久,并且可以反复评分以便得到客观性更高的数据。

就像讨论中所说的,可以采用各种不同的方法来测量想要研究的东西。在下一部分,我们会提供一些可以找到现成测量工具的资源。如果足够幸运地找到了所想要的工具,那么就不需要开发新的工具并确定其心理测量的适当性了。当然,前提假设是测量工具所产生的数据具有心理测量特性。我们之前说过,好的测量工具对于你的研究目的及研究群体来说应该适用性广、可直接引用及产生的数据具有心理测量特性(如信度和效度)。要注意这里说的是“生成数据”,“对于研究目的”及“对于想要评估的群体”。这强调的是一个重要原则:信度和效度是有条件的。它们取决于想要评估的构念的性质、用于估计构念所采用的分数、研究目的及被试群体。两周后 0.60 的再测信度对于情绪状态测量来说足够了,但对于一个比较稳定的品质的测量(如智力)来说是不够的。为了筛选而进行的抑郁测量可能具有让人满意的内容效度,但如果为了诊断或监控治疗,它的内容效度就没有那么好了。一个在评估白人学生的成就时具有较好预测效度的测量工具可能不适合黑人学生。在下一部分,我们将总结一些重要的心理测量特性并提供一些有关评价测量工具的建议和最低标准。本章没有论述测验编制或者测量理论,虽然在研究方案中你可能没有想过这些概念,但是你应该熟悉它们(可参考 Anastasi & Urbina,1997;Cronbach,1990;Cronbach et al.,1972;Guilford,1956;Kaplan & Saccuzzo,2005;Nunnally & Bernstein,1994)。

评价测量分数的信度

无论是在物理科学还是在行为科学中,任何工具的基本标准是它产生的数据是可靠的。多年来,测量学者们定义了不同形式的信度(Kaplan & Saccuzzo,2005),如

重测或时间信度、复本信度、内部一致性等。1972 年,Cronbach 及其同事建议根据概化理论来统一各种类型的信度。本质上就是,各种不同类型的信度以不同的方式归纳数据,它们在理论上和实践上都是不同的。因此,如果只是因为这个工具可靠才选择了它,这在技术上是不正确的,因为是通过测量得到的分数具有某种心理测量特性,而不是工具本身(Thompson,2003)。应该说这些分数具有研究所需要的心理测量特性。然而正确得到这个结论并不容易,这需要你弄清所要研究的事物及打算如何研究它。例如,在研究某一心理构念时,会出现各种心理测量问题,但如果从自然科学的视角进行研究行为,那么问题就没有那么多了(Foster & Cone,1995;Johnston & Pennypacker,1993)。

假设你正在研究某一心理构念。假设是什么?是否假设它在时间上是相对稳定的?在各种不同情境下呢?它是在被试的言语行为中,还是在其运动或生理反应中能够最好地表现出来呢?它是一个支撑并决定被试反应的统一的构念(如焦虑),还是由不同的维度组成(如行为健康)?如果是前者,那么基本测量模型是一个效果指标模型。如果是后者,那么它是一个因果指标模型(Bollen & Lennox,1991)。你应该知道自己的基本测量模型是如何应用到分数的各种信度上的。

在选择测量工具时,对这些及其他问题的回答关系到需要考察的心理测量信息的类型。在理想情况下,能够事先识别出所需要的信息,根据具有这些信息的分数来选择测量工具并提供支持性的信息。至少应该保证测量分数的信度和效度。可以这样说,对于你的研究目的,“需要如下的信度类型。”然后说明都是什么信度,说明你打算采用的分数中存在这些信度。最后,对效度重复同样的过程,这是我们接下来要进行的话题。

评估测量分数的效度

研究中除了希望获得一致性的分数,还希望这些分数具有某些意义,这体现在它们的效度上。一个测量工具可能产生非常可靠的数据但不与任何事物相关。换句话说,测量工具能够产生可靠数据但它们是无效的。在经典测量理论中,反过来是不成立的,有效的数据必定是可靠的。效度是指一个测量工具得到的分数与另一测量工具得到的分数之间的相关程度。这比常见的定义更全面,常见的定义是:如果测量分数能够体现出所要测量的事物那么就是有效的。在本卷中这个定义在更大的程度上意味着它们可能根本做不到这一点,但仍然能够有很高的效度。

就像信度一样,效度也有不同类型。**表面效度**(face validity)指被试能够透过测验项目看出测量目的的程度。**内容效度**(content validity)指测验项目能够代表与所要评估的构念或行为相关的所有内容,忽略无关内容,并平衡各种构念指标的程度(Haynes,Richard & Kubany,1995)。因此,项目选择、类别定义等是关键问题,常常通过专家判断来评估内容效度。

结构效度(construct validity)指测量分数与所要反映的构念的理论基础相关的程度。**聚合效度**(convergent validity)指用于评估相同行为或构念的不同测验得出的测量分数之间的相关程度。如果测量分数与测量其他构念的测量分数不相关,那么就具有**区分效度**(discriminant validity)。**判别效度**(discriminative validity)指在各组

之间出现了预期的平均数差别等。

效标关联效度(criterion-related validity)指测量分数能够反映研究目的的程度。如果测量分数能够预测在其他测验上的分数(即效标),且这个效标常常是实践中的指标,那么这个测验分数就具有效标关联效度。如果效标分数与测量分数同时存在,那么就存在同**时性效度**(concurrent validity)。**预测性效度**(predictive validity)指在测量之后过一段时间才能得到效标分数。

你的测量分数都需要哪些类型的效度呢?为什么这些效度合适呢?或许你的研究就是用来解决这个问题的。如果是这样,那么需要说明哪种类型的效度呢,为什么呢?

需要表达的效度及信度类型取决于研究主题、研究群体、打算采用的特定分数及研究要解决的问题。如果正在研究某一心理构念(如敌意、焦虑、利他主义或沙文主义),那么至少需要说明内容效度和构念效度。例如,如果你研究的是利他主义,需要考察利他主义与其他变量之间的关系。这些预期的关系来自于哪儿呢?正确的答案是与利他主义相关的理论。为了检验这些关系,需要采用能够有目的地测量到利他主义构念的测量工具。这种目的性存在于与测量工具的构想效度有关的研究文献中(参照 Grusec,1991)。

如果研究主题是行为,特别是从自然科学的角度研究行为,那么就没有必要特别关注构想效度。相反,需要关注的是测量工具的内容,特别是那些不便直接研究的行为,如果测量工具中的项目能够代表这些行为,那么就会方便得多。此外,需要知道测量工具产生的行为能够代表在其他时间或地点测量产生的行为的程度。我们之前说过数据能够应用到其他评估场合的程度指的是时间信度。时间信度涉及数据在各种情境下的普遍性(Cone,1978)。例如,如果在实验室观察到夫妇间积极的言语交流,那么这能代表这对夫妇在家里的行为吗?

至于效度,在方法部分只说测量有效是不够的。应该说明对于研究目的来说哪些效度是重要的,并为特定分数中存在这些效度提供证据。如果研究本身就是效度研究那么就更需要这样做了。这是因为陈述进行此项研究的理由需要总结大量与测量工具有关的文献。

任何研究都需要这种详细的总结,你需要知道某个测量工具评估的是不是自己感兴趣的变量及它在你的研究群体中是否具有一致性。如果没有这些重要信息,会给研究结果的解释带来困难。

可以在 4 个地方找到测量分数的心理测量属性:①测量工具的手册;②评估测量工具的章节或综述文章;③评估测量工具心理测量属性的文章;④使用此测量工具的文章。需要注意的是,在使用某测量工具的文章中不会在摘要或前言部分介绍此工具的心理测量属性。因为这不是研究首要关注的,这些数据可能会在方法部分说明。另外,一些不会提及心理测量问题的研究会提供一些与效度有关的结果。例如,反复发现焦虑与抑郁的自我报告测量高度相关会置疑这些分数的区分效度,因为前提假设是这些分数反映的是不同的情感领域。这些结果有助于对这些构念理论的概念重建(见 Barlow,2002;Barlow,Allen & Choate,2004)。

广泛寻找合适的测量工具

我们已经描述了用于操纵研究变量的基本评估方法,也讨论了好的测量工具所具备的重要特征。那么在哪儿能找到这样的工具呢?如果你的研究属于热门领域,那么很可能会采用常用的工具。如果你正在开拓新领域,那么可能不清楚在已有文献中哪些是最好的工具。

一些已出版的汇编提供了评估测量工具的有用信息。其中最著名的是《心理测量年鉴》(*Mental Measurements Yearbook*)(Plake, Impara, & Spies, 2003),由内布拉斯加大学出版社出版,收集了大量不同类型的测验。这本年鉴由奥斯卡·K. 比勒什于1938年首次出版,现在已经是第15版了。许多研究者把这本年鉴简称为*Buros*。除了描述测量工具,超过500名评论者为这些测量工具的心理测量适当性提供了注释。

关于测量的一个相关资源是《纸笔测验Ⅵ》(*Tests in Print VI*)(Murphy, Plake, Impara & Spies, 2002)。与*Buros*不同,《纸笔测验》没有明确回顾测量工具,而是让读者参考心理测量年鉴。但它列出了所有在市场上能够买到的测验,是确定与某一特定测量工具相关的出处和杂志文章的好资源。另外,它提供了与测验目的有关的信息、适宜人群及施测时间。

如果你打算采用行为测量,赫森与贝拉克的行为评估技术词典能够提供帮助(Hersen & Bellack, 1988)。虽然年代有点早,但这本书描述并评价了286个测量工具,是一本很好的入门书籍。200多位研究者评述并注释了最初的测量工具。在本章介绍的图书资源中包括这本书都没有提供测量工具。不要被这本书名中的**行为**一词所迷惑,它包含了许多被人格心理学家和临床心理学家所熟悉的传统特征取向的评估工具。

斯威特兰和凯泽在1983年出版了有关测验参考文献的丛书(Sweetland & Keyser, 1983),并在1983年与1994年间出版了10卷。他们索引并描述了超过3 500个测量工具。分成商业、教育及心理学3个主题。可以从中找到想要的测量工具,并通过相关参考文献来了解更多信息。本丛书提供了很多关于信度、效度及施测程序的信息。

斯威特兰和凯泽(Sweetland & Keyser, 1997)也简要介绍了根据主题排列的且在市场上能够买到的测验,包含700多个。每一个都介绍了测验目的、测验的适当性、成本、可得到性及出版者。如果考察患者群体或常见变量,《精神病评定量表手册》(*Handbook for Psychiatric Rating Scales*)(研究与教育协会,1981)值得一看。这个手册提供了每个评定量表的信度和效度、描述性信息、适用人群及相关参考文献。其他有关特定量表的书籍有 Fisher and Corcoran(2000a, 2000b); Robinson, Shaver, and Wrightsman(1991)。测量焦虑时可参考 Antony, Orsillo, and Roemer(2001)。Nezu, Ronan, Meadows 和 McClure(2000)提供了大量用于测量抑郁且得到实证支持的工具。如果研究儿童的在校行为,那么可参考 Kelly, Reitman 和 Noell(2002)。Buros 测验中心的网址(http://www.unl.edu/buros)是个不错的网上资源。**测验评述在线**

(Test Reviews Online)评述了出现在《心理测量年鉴》丛书中的测量工具,且每月更新一次,不过是收费的。测验工具有 2 000 多种,用户可以下载相关测验的评述及细节描述。注意,许多高校图书馆都有《心理测量年鉴》数据库的 CD 版。CD 里包含《心理测量年鉴》中最新的测量工具,每 6 个月更新一次。

与行为医学有关的较好资源是健康与社会心理测量工具(HaPI™)数据库。HaPI™提供了 8 000 多个没有发表的测量工具,每季更新一次。已订阅此数据库的图书馆及个人可获得 CD 版。多数测量工具与行为医学有关,可评估疼痛、生活质量及药物疗效等变量。HaPI™也包含了医学相关学科的测量工具,如心理学、社会工作、职业疗法、物理治疗及言语和语言治疗。最后,HaPI™提供了测量工具的参考文献,用户可以通过行为测量数据库服务(BMDS)中的工具传递服务来订购。

另一个较好的资源是教育测验服务数据库服务(ETS),但在市场上买不到。据说 ETS 是世界上最大的数据库,包含从 19 世纪 90 年代到现今的 20 000 个测验及其他测量工具。

最后,如果这些资源都帮不上忙,那么可以考虑 PsycINFO。用描述变量的术语或同义词在这个数据库中搜索,至少可以找到相关文章。如果幸运的话,某些参考文献可以提供一些测量变量的方法。古德曼和米切尔(Goldman & Mitchell, 2003)出版了一个词典,包括未出版的、在市场上买不到的并且在大约 36 个好杂志上出现过的测量工具。你可以在这里选择研究所需要的、带有心理测量信息的测量工具,而且使用是不收费的。

如果手头没有重要的心理测量信息要知道应该做什么

如果选择的测量工具不合适,并且没有其他工具可选择,应该怎么办呢?如果明白了之前的内容,那么应该知道需要报告测量工具的心理测量特征。现在让我们来看一看更重要的心理测量特征,以及当测量工具不具备这些特征时该怎样做。

评分者信度

可靠的记分工具对于任何研究来说都是必要的。即使已有研究证明你所使用的工具记分是可靠的,但仍希望证明它在你的特定研究中是可靠的。我们之前提过这一点,在此再强调一下。如果采用的是自我报告方法,另一个人对答案的 25% 进行独立评分;如果两部分的分数不一致,看看问题出在哪儿并改正它。如果采用的是结构性访谈,要尽可能全程记录访谈,并让评分者对录音带或转录后的文字进行独立评分。我们在第 7 章说过,至少 20% 的评分应该是一致的。如果 20% ~ 25% 的评分不一致,那么就需要进行干预了,可能其中的一个或者两个评分员不认真。要让评分员保持或提高正确率,并且之前的所有评分(不只是信度数据)都需要重新进行。

如果等级评分由他人完成,需要比较两个评分者对同一被试 20% ~ 25% 的回答上的评分等级。可以通过计算一个简单的相关来完成,或者采用某种百分比一致性方法来得到更精确的信息。例如,可以把在某一项目上的一致性定义为两个评分

者在这个项目上的评分等级相同。或者可以要求得宽松些,允许它们在彼此的一个评分等级内就可以了。用一致性的数目除以总的评分等级数目就是一致性百分比。用 kappa 系数可纠正随机一致性(von Eye & Mun,2005)。确保信度中的准确率水平至少与分析中分数的准确率相同。例如,根据在“社会线索意识”测验中的总分进行分组,需要保证不同评分者产生的总分相似。可能在个别项目水平上的评分不同,但被试分组应该是相同的。

如果采用的是直接观察编码系统,可以考虑采用录像或录音带来永久记录被试的行为。观察者可以对此进行独立编码,并比较他们的数据。同样的,可以随机选择观察百分比,并计算一致性。如果不能采用录像或录音记录,最好让观察者成对工作,并检查他们的一致性。让主要观察者随时待命,这样可以保证相同的两个人并不总是一起工作,以避免一起工作的观察者在定义行为时的互相影响。如果不能采取这些看似严密的防范措施,那么可以只使用一个观察者,并让另一位观察者定期检查他的工作。然而,这不是理想的安排,因为检查者的出现会影响到初级观察者的行为(Romanczyk et al. ,1973)。如果采用这种方式,要训练观察者使其达到高水平的一致性,这样即使他们在评估时做得更好,也不会过多影响到结果。

即使数据来自于档案(如医院中的病人档案或县法院的记录),也不能忽视评分者信度。你可能认为通过当前日期与第一次住院日期之差来决定慢性精神分裂症是一种惯例,然而第一次住院的日期可能有多个,这取决于你如何看待它。此外,你或你的助手在记录住院日期或对日期进行相减时可能会出错。为了准确起见,在生成数据时记得时常检查这些步骤。保证可靠评分的一个关键因素是对评分者和观察者的训练。在研究正式开始之前,可让评分者或观察者在模拟数据上练习。在确保他们能够独立工作,并且在练习数据上有很好的一致性后再让他们对真实数据进行评分。最好多进行训练,使观察者的一致性比你所能接受的高出 10% ~15%,这是因为当评分者和观察者在收集“真正”数据时一致性常常会下降(Romanczyk et al. ,1973)。

时间信度

如果你的研究需要测量分数的时间信度,而你却不知道这一信度,可以采取以下措施。在正式研究开始之前可在合适的时间内重复施测,然后计算被试群体在这两次评估中测验分数的相关程度。换句话说,就是自己建立分数的时间信度。另一种方法是,通过研究设计来控制因为时间流逝而可能发生的分数变化。例如,增加一个控制组,或者采用一个更复杂的、包含控制组和其他区组(如所罗门四组)的设计。还可以在引入自变量的前后重复评估被试。在被试内方法中,每个被试都是自己的控制组,控制练习或时间流逝导致的分数变化。这种方法可以知道重复评估的个体或群体中是否有绝对的分数变化。如在本章前面所述,即使相关分析表明信度很高,平均数也可能发生变化。

项目信度

内部一致性评估是建立新测量工具的最重要特征之一,所以首先要掌握内部一

致性的信息。如果没有,可以根据已有的标准数据进行计算(如测验手册)。可采用库德-理查森公式21(Kaplan & Saccuzzo,2005;Nunnally & Bernstein,1994)来计算内部一致性信度。如果采用内部一致性低的量表进行测量,会很难解释数据,这是因为内部一致性低的工具测量的是多个事物,相同的分数可能意味着不同的事物。

不能在进行研究的同时计算内部一致性信度。如果内部一致性信度低,再来修改工具就来不及了。提高内部一致性,需要以新的方式来组织测量工具中的各元素,从根本上改变工具的结构并放弃已有的标准和效度信息。如果起初没有有关内部一致性的数据并且想使用这个工具,我们需要测量足够大的被试样本,并对此数据进行因素分析,计算被试在每个因素上的分数。可以把在理论和实践上与你的构念最符合的因素作为自变量。当然,这些新开发的量表可能没有标准或效度信息。此外,合理的因素分析需要足够大的样本,这在第7章说过。这取决于量表中的项目数,可能最少需要100人,如果重复因素结构,人数需要增加一倍。当然,如果分析中包含更多的项目或测量工具,也要增加人数。一个更好的办法是在正式研究开始之前收集小样本数据并分析内部一致性信度。

上述讨论多数适用于自我报告等级评分及评分者等级评分方法,但至少在概念上适用于所有评估方法。由于篇幅所限,对此不再多说。无论采用何种方法,不管研究主题是心理构念还是实际行为本身,评估工具的一致性都是重要的。例如,假设你正在直接观察鸟蛛的捕食行为,你的假设是你所编码的几个特定反应是同一捕食行为类别中的所有反应。如果一些涉及筑巢,一些涉及觅食,一些涉及生殖,那么一个单一的分数如何能代表所有的行为呢?你又如何表示呢?

聚合及区分效度

前文提过聚合及区分效度,但没有详细讨论。这在采用自我报告方法测量心理构念时特别重要,因为自我报告测量常常缺少区分效度。也就是说,它们与已有测量普遍个体差异变量的工具有很大程度上的相关。这些普遍的变量包括智力、社会经济地位、教育水平及说出自己的社会期望的倾向。生成新工具时一定要说明它们不仅仅是测量这些变量的另一种方法(Campbell,1960)。这是因为社会期望的渗透使得新的自我报告测量必须独立于社会期望(Edwards,1970)。在文献中已有很多测量社会期望的好工具,所以我们不需要再开发一个。要确保你所使用的量表与其他社会期望量表无关,或只有微弱的关系。与此同时,需要注意的是评估社会期望的各种方法(如 Crowne & Marlowe,1960; Edwards,1957; Wiggins,1959)是不可相互替换的(Edwards,1990; Helmes & Holden,2003)。保罗赫斯(Paulhus,2002)认为造成自我报告曲解的原因有3个:印象管理,为提高自身积极品质而进行的自我欺骗,为否定自身消极品质而进行的自我欺骗。三者与人格量表的相关程度不同(Paulhus,2002)。爱德华的观点在自我报告测量中具有普遍影响。因此,要确保你的量表至少与这种类型的社会期望无关。在比较自己与他人的量表时还存在其他类型的反应定势(如渴望赞同;Crowne & Marlowe,1960),但最普遍的就是社会期望。

如果测量工具不能独立于社会期望,那么需要在研究中解决这个问题。第一种

方法是减少那些能够增加自我报告偏向的特征要求(如匿名回答;更多的方法见Paulhus,1991)。另一种方法是同时采用社会期望量表和其他量表,并亲自计算它们之间的相关。确定哪种形式的社会期望对自变量的影响最大。如果相关不显著,那么就可以排除社会期望对结果的影响。如果存在相关,首先需要确定前人理论和研究中的相关程度是多大。例如,抑郁个体倾向于以相反的方向曲解信息(Beck, Rush, Shaw & Emery, 1979),因此自我报告的抑郁与社会期望之间可能存在负相关。当然,如果相关很高那么就要停止使用此量表。如果一种或多种社会期望不应该与你的测量工具存在相关,你希望在统计上控制社会期望的影响,可以采用偏相关或协方差分析,也可以考虑滤掉社会期望分数比较高的被试数据。

聚合效度也很重要。坎贝尔和菲斯克(Campbell & Fiske, 1959)早在多年前就提倡计算测量工具之间的聚合性。需要注意的是任何工具的测量分数都部分地体现出工具本身的特征,作者们建议采用多特质-多方法矩阵(Multitrait-multimethod matrixes)来找出这种方法差异,并在此过程中建立聚合效度和区分效度,要知道,怎么强调测量构念时聚合效度的重要性都不过分。我们的语言结构在概念间的关系之中自动生成,当我们采用基于语言的评估工具时,构念之间至少因为语言结构而互相关联。某一测量工具的效度不应该只包含常用测量工具间的相关关系。如果在某一测量工具的分数下真的存在一个构念,那么评估它的方式不止一种。原因在于,如果某个测量工具能够完全操纵一个概念,那么这个概念就不再是假想上的了。因为心理特质的一个基本特征在于它的假想性,如果不具有假想性,那么它就不是一个特质。测量工具间的关系能够建立聚合效度并同时扩展潜在心理变量的结构效度的基础是你所测量的概念具有假想性。

如果对于你的测量分数或被试群体来说没有现成的聚合效度,那么需要在你的研究中建立聚合效度。最简单的方法是采用其他方法来评估变量并在数据收集完毕后进行比较。如果不同的方法之间存在相关,在数据分析时可以互换使用或者两个都用。如果不存在相关,需要分别处理,并在讨论部分进行说明。

采用其他的评估方法并不意味着要进行另一个自我报告测量,又一个评分者等级评定,等等。例如,自我报告数据可作为评分者等级评定或直接观察数据的补充,有时补充数据非常有用。例如,对饮酒进行自我报告是研究方法之一,可以把自我报告与观察者如配偶或室友或朋友的报告进行比较;或者采用隐蔽的方法如计算垃圾堆里空啤酒罐或液体瓶子的数目(Webb, Campbell, Schwartz & Sechrest, 1966)。一些行为会产生直接的产物或痕迹,可作为佐证。自我报告的作业完成情况或花在论文或博士论文上的时间,正确解决的问题数目或写出的文章页数都可作为其口头报告的佐证。

跨情境的概括化

在评估一个人的行为时,需要特别关注在一个情境下收集到的数据能否应用到其他情境,多数情况下需要研究者亲自考察这个问题。也可以采用研究者们最常用的一句警告,如"在解释这些结果时应该谨慎,因为不知道被试在家(教室,邻里等)的表现是否类似。"如果你的研究是应用型的,那么大概会在最感兴趣的情境下实

施研究，但仍未能解决情境一般化的问题。如果研究的是心理构念，跨情境一般性则常常是理论上的，很少需要真正检验。这是个值得思考的问题，特别是当你关注研究的生态学效度时。

谨慎修改他人的测量工具

你是否希望修改他人的测量工具来研究自己的问题？例如，你发现一个很好的用于评估自信的量表，并想修改它用于多重人格障碍的个体，让被试根据每一种人格填写这份量表。当然，这意味着需要改变这个工具的指导语和重点。

它的原始心理测量属性还有用吗？你不能确定。从逻辑上说，对工具改变得越少，保留的原有属性就越多。例如，某个明尼苏达多相人格问卷—2（MMPI—2）分量表产生的分数会与从整个 MMPI—2 中选取相同项目产生的分数相同。然而，除非收集一些数据来验证，不然仍不确定。一种方法是进行预实验。在简短版与原版 MMPI—2 上的分数会存在相关（就像评估复本信度）。并且，如前所述，采用其他测量工具有助于评估改变后的工具效度是否可以直接应用到研究项目中。

在评价和选择测量工具时避免常见错误

描述测量工具时普遍存在几种常见错误，我们要避免这些错误。首先，不要认为测量工具的名字能够代表它的功能，许多新手甚至有经验的研究者都陷入了这个误区。这并不一定如此。仔细检查测量工具的内容：测量工具的名字可能与所要评估的构念不匹配。这个错误在心理学中被称为**叮铃谬误**（jingle fallacy）（Kelley，1927）。

特别是对于通过因素分析而产生分量表的问卷来说，名字通常不能够反映问卷的内容。简单来说，因素分析只关注数据，找出哪些因素之间紧密联系，并指出其在数据上的关系；研究者们根据项目之间联系的紧密程度来决定哪些项目属于一个因素。项目的内容与因素分析中的数学计算没有任何关系，这意味着同一因素内的各项目可能反映了不同内容。然后研究者必须想出一个看似适合项目的名字。除非测验的开发者进行了验证性的因素分析来检验在概念上相关的项目是否紧密相关，不然名字只是研究者事后的命名。所以，不要认为名字如实反映了量表内容，也不要认为量表充分反映了你想要研究的构念。

另外，不要认为显著相关就是高相关。许多量表使用者都这么认为。测量同一构念的两个测量工具之间的显著相关并不意味着你想要考察的测量工具具有效度。重要的是相关的大小，而不是它的统计显著性。劳动生产率的自我报告测量与机械化评估的相关程度为 0.02，且在 $p < 0.05$ 水平上显著，但它们的共同方差只有 4%。这真的意味着二者可以互换吗？

并且不要因为一个自我报告测量与其他自我报告测量相关就认为它能够评估行为。如前所述，坎贝尔和菲斯克（Campbell & Fiske，1959）很早就说过，采用相同方法来评估同一构念的两个测量工具之间的相关并不能“证明”新的测量工具有

效。方法方差(如,测量分数中的某些变异源于测量工具本身)能够解释结果。当两个自我报告测量采用相同方法时,它们的分数可能会因为相同的评估方法而相关。检验间接测量(如自我报告)的真正方法是看它是否与评估相同现象的更直接的测量方式相关。总的来说,评估同一事物的不同方法之间的相关程度应高于采于相同方法来评估不同事物之间的相关。

最后,不要采用单一项目测量,因为采用单一项目来测量构念或行为非常不可信。然而,新手们常常认为某个被试对一个问题的回答,如"你的爸爸或妈妈是否酗酒?"是判断其是否成长在酗酒家庭中的一个好方法。这是一个问题,特别是在测量被试特征,如描述或分类(例如,作为一个自变量)的时候非常常见。同样,新问卷的设计者常常通过单个项目上的回答来解释他们的工具。除非你能够说明单个项目是可信且有效的,否则不要这样做。

获得测量工具复本

检验测量工具的内容首先需要获得此工具的复本,这非常困难。杂志文章很少发表测量工具的复本,一些在市场上能够买到,但是很多都买不到。

可以联系工具开发者并向其索取测量工具复本。确保找到开发者现在的地址。专业目录如《美国心理学协会会员注册》(*American Psychological Association Membership Register*)及其他专业组织的成员列表(如儿童发展研究协会及美国心理会)会定期更新,应该会帮助你找到某个人。通常最简单的方式是在 Google 上搜索这个人。找到这个人时,问问他/她这个工具是否有版权。如果有,必须获得使用这个工具的书面许可。

如果找不到这个作者,或者作者没有给你答复,或者你想从多方面下手,还可以从已发表的文献中寻找其他使用过这个工具的人,这个人一定有测量工具的复本。也可以问问你的导师谁有这个工具复本,然后找到这个使用者并索取。可以使用你的导师的名字(在允许的情况下)以提高获得回复的几率。记住,即使很费劲地得到了工具复本,但从中获得的有关其重要心理测量属性的信息可能很少。

补充资源

一般测量教科书及文章

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Butcher, J. N., Graham, J. R., Haynes, S. N., & Nelson, L. D. (Eds.). (1995). Special issue: Methodological issues in psychological assessment research. *Psychological Assessment*, 7, 227-413.

Campbell, D. T., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: HarperCollins.

Kaplan, R. M., & Saccuzzo, D. P. (2005). *Psychological testing: Principles, applications, and issues* (6th ed.).

Belmont, CA: Wadsworth.

Nunnally, N. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Thompson, B. (Ed.). (2003). *Score reliability: Contemporary thinking on reliability issues.* Thousand Oaks, CA: Sage.

问卷方法(自我报告,他人评估)

DeVellis, R. F. (2003). *Scale development: Theories and applications* (2nd ed.). Thousand Oaks, CA: Sage.

Edwards, A. L. (1970). *The measurement of traits by scales and inventories.* New York: Holt, Rinehart & Winston.

Fink, A. (2002). *The survey handbook* (2nd ed.). Thousand Oaks, CA: Sage.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation*, 22, 127-160.

Stone, A. A., Turkan, J. S., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S., & Cain, V. S. (Eds.). (2000). *The science of self-report: Implications for research and practice.* Mahwah, NJ: Erlbaum.

直接观察

Bakeman, R., & Gottman, J. M. (1986). *Observing interaction: An introduction to sequential analysis.* Cambridge, England: Cambridge University Press.

Cone, J. D. (1999). Observational assessment: Measure development and research issues. In P. C. Kendall, J. N. Butcher, & G. N. Holmbeck (Eds.). *Handbook of research methods in clinical psychology* (2nd ed., pp. 183-223). New York: Wiley.

Suen, H. K., & Ary, D. (1989). *Analyzing quantitative behavioral observation data.* Hillsdale, NJ: Erlbaum.

von Eye, A., & Mun, E. Y. (2005). *Analyzing rater agreement: Manifest variable methods.* Mahwah, NJ: Erlbaum.

隐蔽测量

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally.

自我观察

Christensen, T. C., Barrett, L. F., Bliss-Moreau, E., Lebo, K., & Kaschub, C. (2003). A practical guide to experience-sampling procedures. *Journal of Happiness Studies*, 4, 53-78.

Cone, J. D. (Ed.). (1999). Special section: Clinical assessment applications of self-monitoring. *Psychological Assessment*, 11, 411-498.

Foster, S. L., Laverty-Finch, C., Gizzo, D., & Osantowski, J. (1999). Practical issues in self-observation. *Psychological Assessment*, 11, 426-438.

Stone, A. A., Kessler, R. C., & Haythornthwaite, J. A. (1991). Measuring daily events and experiences: Decisions for the researcher. *Journal of Personality*, 59, 575-607.

思考与练习

选择适当的测量方式

☐操作变量

☐评估工具的信度

—评分者信度

—重测信度(时间信度)

—复本信度

—内部一致性(项目信度)

—确定研究中需要哪种类型的信度

☐评估工具的效度

—表面效度

—内容效度

—效标关联效度

· 同时效度

· 预测效度

—结构效度

· 聚合效度

· 区分效度

· 判别效度

—确定研究中需要哪种类型的效度

☐寻找合适的测量工具

—通过测量工具汇编

—通过电脑数据库

—通过已发表的文献

☐没有关键心理测量信息时的做法

—评分者信度

—时间信度

—项目信度(内部一致性)

—聚合及区分效度

—跨情境的一般化

☐慎重修改他人的测量工具

☐获得测量工具复本

9 选择合适的统计方法

没有什么比**统计**这个词更让初级研究者恐惧的了,即使能够轻松通过严格的研究生统计课程,但让他们为自己的研究选择合适的统计方法时,他们却表现出对统计的无知。

为什么即使最好的学生有时也会在统计方面感到困难?讲授统计课程所使用的方法是产生问题的一部分原因。统计学教师通常采用下列方法中的一种:①数学的;②操作步骤。传统数学方法讲授的是推导公式及统计分析的概念基础,操作步骤方法讲授的是如何进行统计计算。根据我们的经验,这两种方法都没有讲述如何根据研究问题和研究假设来选择合适的统计方法。问题的另一部分原因在于"统计焦虑",源于早期对一个人数学能力的否定或在统计课上的失败经历(多数是因为老师教得太差了,而不是学生的能力问题)。数学焦虑,反过来会导致学生回避使用数字并因此失去在这个领域培养自信的机会。回避及缺少练习导致学生依赖于他人来选择合适的统计方法,结果,他们从来都没有学会在这个领域独立解决问题。

本章将帮助同学们挑选出哪些统计方法最适用于你们的研究设计。我们强烈建议你们在完成论文提案之前阅读本章并照着做,而不是在收集数据之后才考虑令人恐惧的统计问题。这就是把本章放在数据收集之前的原因。导师常常会面临一项艰难的任务,那就是给那些收集完数据后却不知道怎么做的学生一些建议,这是因为他们在选择测量工具和进行研究设计时没有考虑统计方法。

我们假设读者已经学习过了统计课程,并且至少熟悉不同统计方法的名字及它们的一般用途。因此,本章不会讲述如何计算、导出或编写特定的分析。相反,我们会主要考察与两组或多组个体之间的差别及在某一组个体中与变量间的关系有关的统计方法。对于我们所讨论的大多数策略来说,我们假设你的研究包含自变量(或预测)和因变量(或效标)。本章末尾会简要讨论模型检验分析,模型检验分析可用于考察一组变量的潜在结构或同时考察变量间的关系。

本章没有包含现今所有的统计检验方法,没有提及分支学科中的专业方法,省略了最常用方法中的序列分析及其他时间序列方法,元分析技术及计算效应大小的特定统计方法,复杂的纵向分析(如层次线性模型),多层次模型及生存分析(survival analysis)。我们也没有包含存在争论的统计方法。相反,我们讨论了广泛使用的统计方法及它们在分

析计划中的潜在作用。对于资深的统计学家来说,本章看起来似乎太过简单,而对于缺少经验的人,本章看起来似乎又过于复杂。

本章的目的在于选择合适的统计方法。一旦选择了正确的统计方法,就需要参考其他资料来确保自己理解这些统计方法并确保对你的研究来说它们是最恰当的。在统计分析前都需要准备什么呢?

早点加强统计知识

如果你刚刚开始研究生涯，要多参加研究并主动进行数据分析，然后与教师或研究主任一起工作。问问为什么某一特定的分析方法比其他方法更合适。主动负责研究的统计工作会强迫自己去学习与所使用的统计相关的参数和选项，这也有助于你采用系统脱敏方法来治疗统计焦虑！

另一种方法是看用母语写的统计书，多数统计书都包含所有的统计专题。另外，多数教授都会概要描述一个统计方法以便同学们使用。书籍与网站信息在复杂性、重点、数学与解释性材料之间的平衡、覆盖深度及（有时）术语和符号系统上会有所不同。因此，有些材料你会用得很方便，有些则不是，要寻找自己能够看懂的书。本章中会提及一些有用的统计书籍。在本章末尾我们也会提供一些额外的资源。另外，一些主要的计算机程序手册能够提供很好的统计信息。重点在于寻找那些以自己能够理解的方式来解释统计方法的资源，即便是本书的两位作者有时也会在哪些统计书最容易理解这个问题上意见不一致。

首先考察自己的研究：建立一个分析计划

选择合适统计方法的第一步是确定分析的目的。最初应该清楚地指出你的统计想要回答的问题，逻辑起点在于研究的问题和假设。不过应该超越这个逻辑起点，因为你所采用的统计方法不仅仅是为了检验假设。例如，你可能想要考察各种因变量之间的关系，或者检查三组实验被试在人口学特征上有没有差别。在下面部分，我们会逐步地指出符合特定需要的各种统计方法。我们建议读者遵循分析计划的每一步。

列出想要分析解决的问题或目标，然后列出回答每个研究问题所需要分析的所有变量，并检验每个假设。例如，如果问题是，“样本的人口学特征是什么？”列出想要考察的人口学数据，接下来思考如何计算被试在每个变量上的得分，最终形成一个问题表。在每个问题下面列出用于分析的分数。

注意我们说的是“分数（scores）”而不是“测量工具（measures）”。例如，只说”MMPI—2”是不够的，明尼苏达多相人格问卷—2 会产生很多分数。当然不必考察每个分数，只需要列出想要分析的分数就可以了。

因而需要澄清每种测量方法是如何产生数据的。可以复印一份测量工具，生成几个被试的模拟数据，对数据进行计分，然后进行分析。每个工具会产生多少分数？需要分析哪些分数来考察研究假设和问题？把这些放到变量列表中。在这之后，需要列出想要回答的问题及需要分析的变量。

我们喜欢把分析计划分成初步分析部分和假设检验分析部分。在初步分析部分需要：①考察数据的特征；②考察测量工具的属性（如信度，效度）；③检查可能存在的混淆变量；④考查统计假设。假设检验分析可用于考察数据是否支持预测。此时，列出想要考察的问题/假设，及用于考察每个假设的变量。然后写下用于回答问题或检验假设的特定分析方法。表 9.1 是科利特 · 洛德（Colette Lord，2005）博士论文中的分析计划部分，考察的是间接攻击与感受到长辈的社会支持之间的关系。

其次,选择一个想要回答的问题。指定用于回答这个问题的特定自变量及因变量(分数)。例如,如果问题是:"注意缺陷多动障碍(ADHD)及非 ADHD 男孩与女孩在回忆社会材料上是否有差别?"你会把诊断(ADHD 或非 ADHD)及性别(男孩或女孩)作为自变量,并把回忆分数作为因变量。然后问问自己:通过比较两组被试在回忆分数上的差别能否解决问题?或者一组被试在不同变量上的分数之间的相关能否解决问题?对于上面的例子,第一个问题的答案可能是肯定的:需要比较男孩与女孩之间的差别,ADHD 与非 ADHD 组之间的差别,等等。这需要通过一些统计方法来比较各组之间的差别。用于比较各组差别的常用方法包括组间、组内、混合及准实验设计。组比较的统计方法适用于比较不同组被试(如男人或女人)在所得分数上的差别,或者同组被试在不同时间点或不同条件(如处理前或处理后)下所得分数的差别。

需要进行组比较的研究设计把不同组的被试作为自变量。每一组是一个**水平**(level)并有一个名字。在上述例子中,性别有两个水平(男孩和女孩),因此需要分别分析男孩和女孩并进行比较。诊断也有两个水平(ADHD 与非 ADHD)。如果研究者对性别不感兴趣,就没有必要把性别设为自变量,也就没有必要在分析中考察性别差异。

如果想把所有被试作为一组并考察分数之间的关系,那么可考虑相关设计。例如,如果研究问题是,"健康信念(通过在一个多项问卷上的分数来评估),社会支持(多项问卷上的分数),及坚持药物疗法(定期血液水平测定的平均值)之间的关系是什么?"这个研究中所有被试属于一组,没有把被试分成多组(水平)。被试在社会支持及健康信念测量上的分数与他们的坚持值相关。在相关设计中,变量可以是连续的或是分散的,虽然多数是连续的。我们接下来会探讨用于考察关系型问题的相关统计。

专栏 9.1

博士论文分析计划样本:老年退休社区居民中的间接攻击和社会支持

初级分析:

1. 处理缺失数据
   —在自我报告问卷中用平均数代替缺失项目值
   —如果有大量缺失数据可寻找其形态
2. 考察筛选出的分数/数据的分布
   —看分数范围,最小值及最大值
   —如果某些东西看起来失衡,看看数据是否存在错误
   —找出非正态分布的变量
   —找出潜在异常值
3. 考察所有分数的信度
   —计算直接和间接攻击分数的 α 值
   —删除会减少量表信度的项目,重新进行信度分析
   —重复这一步骤,直到对量表信度满意

专栏9.1 （续）

—修改能够反映这些变化的量表生成语法
—计算社会支持总分及分量表分数的 α 值
4. 考察潜在协变量/控制变量
—以所有攻击和支持分量表分数作为因变量进行 ANOVAs；自变量：性别（2 个水平），婚姻状况（2 个水平），退休社区成员（6 个水平），所使用的餐厅（2 个水平），恋爱关系（2 个水平），与子女的联系（2 个水平），独自生活（2 个水平），生活在独立住宅中（2 个水平）
—考察所有分量表攻击和支持分数之间的相关：教育水平，年龄，生活在退休社区的时间，现在及退休前的收入，在进行同龄人提名时不太了解其他人的人数
假设检验：
假设 1：攻击量表的聚合效度和区分效度。
—考察不同方法下攻击分量表间的双变量相关形态
假设 2：自我与同龄人对间接攻击的报告会预测对已有情感支持和来自于退休社区的情感支持的整体感知。
—分层多元回归—因变量/效标变量：
社会支持总分，区组 1：输入协变量，区组 2：输入自我报告间接攻击分数，同龄人报告间接攻击分数
—分层多元回归—因变量/效标变量社会支持退休团体分量表分数，区组 1 输入协变量，区组 2 输入自我报告间接攻击分数，同龄人报告间接攻击分数。
—检查是否违反回归假设
—正态：评估残差是否正态分布
—考察分布直方图的频率
—方差齐性：考察残差的正态概率图
—考察预期残差和实际残差的关系图
—考察标准残差与由回归分析产生的标准预期值的散点图
—线性：画一个有关标准残差与由回归分析产生的标准预期值的图
—多重共线性：考察预期值与容许残留值之间的相关

注：这是 Colette lord 的 *Indirect Aggression and Social Support Among Elderly Retirement Community Residents*（Lord，2005）这篇论文中的部分分析计划。洛德把间接与直接攻击量表应用到老年人退休社区群体中，考察其信度和效度，及间接攻击与社会支持之间的相关。她在退休社区中单独的房间内收集数据。她的实际分析计划列出了 SPSS 命令，特定变量的名字及用于检验额外假设和考察探索性问题的分析。这个分析计划的使用经过了作者同意。

一些复杂的相关统计都与模型检验有关。模型检验包括结构方程模型和线性增长模型。这些技术与其他类型的统计分析的不同之处在于它们检验假设的方式。许多组比较和相关统计通过拒绝虚无假设及与零在统计上显著不同来检验假设。相反，模型检验考察的是实际数据适合用于描述变量间关系的模型（理论）的程度。在这里统计反映的是数据适合用于支持你的理论模型的程度。

一些研究主题可采用组比较或相关设计。如果想知道高智商的儿童是否比低智商的儿童在与同伴交往时存在更多的困难，可以采取两种方式进行研究：①收集 IQ 分数和同伴交往分数并考察二者之间的关系（即计算二者的相关）；②根据 IQ 分数把儿童分成两组或多组（如高 IQ 与一般 IQ），并采用组比较的统计方法来考察各

组在同伴交往分数上的差别。

如果不确定采用哪种方法,可以再看看提案中的研究设计部分(见本书第7章),弄清楚自己想做什么。当自变量是自然连续体(如智力)或者考察能够预测因变量的最佳自变量组合时可采用相关设计。一些研究者认为当不需要操纵自变量时,相关设计比组比较更适合。另外,当变量间存在线性关系时,不采用相关统计方法,却把连续变量分成两组或三组来进行组比较,这会降低统计检验力(Cohen,1983)。当自变量是自然分类(如种族、性别)或自变量是以非线性方式进行的变量组合(如精神病诊断,就是基于几种行为的出现、缺失或程度)时,组比较设计比相关设计更适合,操纵自变量时组比较也更适合。

第二步是弄清楚哪些因变量适合参数检验。参数检验包括方差分析、多元回归、皮尔逊 r 相关,及 t 检验。参数检验的假设是样本群体中分数的潜在分布是正态的(特定的参数检验也有额外的要求)。一个简单的但也很通俗的经验法则是序数或分类数据不符合这个假设,因此应该采用非参数方法来分析。然而这个经验法则也有例外(如 Myers,DiCecco,White & Borden,1982)。连续数据比分类数据更符合参数统计的要求。因此,处理连续数据时首先考虑参数分析。如果数据特征不符合参数分析的假设,可考虑非参数检验(如用卡方检验来代替 t 检验或方差分析)。这特别适用于没有稳定地违反特定假设并且不能够避免违反假设的情况。

参数和非参数分析的选择常常取决于**因变量**的性质。逐个考察每个因变量。并不是所有的变量都具有相同的特征。例如,一个常见错误就是认为可以采用相同的分析来检验两个或多个实验组是否在年龄、种族、性别、教育水平及其他人口学变量上相同。一般采用参数统计来考察年龄和教育水平,采用非参数统计来考察性别和种族变量,这是因为它们是分类变量。同样,可以采用完全不同的统计方法来回答研究中的不同问题。不要认为一个检验能够做所有的事情!

此时,把需要用统计来回答的问题列出来,同时也需要列出每个问题下的自变量和因变量。对于每个因变量,要确定是参数检验还是非参数检验更合适。最后,当把被试分组并进行组比较时(组比较统计),确定是采用某一组在一个或多个测量工具上的分数的相关(相关统计),还是检验同时包含多个自变量和因变量的某一特定模型。接下来,我们会依次介绍组比较统计(参数和非参数)和相关统计(参数和非参数),也会简要介绍一些常见的模型检验分析。这有助读者把合适的统计名称添加到分析计划中去。

组比较统计方法:参数统计

方差分析和 t 检验是比较常见的用于考察各组差别的参数统计。如果确定某一种方法适用于你的分析,还需要从它的各种变式中进行选择。

方差分析和 t 检验

为了选定某一参数组比较统计方法,首先要明确在回答某个特定研究问题时需要多少自变量,其次要明确每个自变量有多少水平(条件或区组)。把每个自变量

分成独立的组(如,条件:回忆启动、再认启动、无启动[3 个水平];时间:前、后、随后[3 个水平])。最后,对于每个自变量,是否需要对包含不同组被试的各个水平进行比较或比较不同条件下**同一组**被试在同一变量上的差别(如,不同时间点呈现不同实验刺激;自变量的时间类型)? 回答了上述问题,就可以利用流程图 9.1 来选出最合适的统计方法。

照着图 9.1 做,你会做出最合适的初步选择。这个流程图会引导你考察你的数据是否符合统计假设。如果不符合,确定这个统计方法是否违反了假设,或者是否有转换数据的方法,或者调整统计方法来减轻这一问题带来的影响。许多学生害怕使用数据转换,但是进行对数、平方根或其他类型的转换来迎合非稳健统计假设并不是欺骗。事实上,这可能有助于避免犯 I 类或 II 类错误。

如果想不出办法来弥补所违反的假设,流程图会引导你选择其他需要较少假设的统计方法。当然,假设越少,检验力越小。通常情况下,参数统计比非参数统计的检验力更大,所以最好采用参数检验。马克斯韦尔和德莱尼(Maxwell & Delaney, 2004)指出,这条经验和多数经验法则一样过于简单化:在某些情况下,非参数方法会比参数方法的检验力更大。哪种检验力更大取决于数据集的特征。如果不知道结果会怎么样并且自己又是个新手,那就需要考虑得更实际些:选择你或你的统计顾问更熟悉的统计方法。

举个例子来回顾一下这些步骤。初步考查数据时,想知道男人和女人在管理类型问卷上的回答是否有差别(分数范围为 30 ~ 150)。自变量(性别)有两个水平(男人和女人)。因为要进行组比较,所以可以采用 t 检验或 ANOVA。由于只有两组,所以 t 检验是最合适的。此外,由于两组被试不同,所以适合独立样本 t 检验。当只有两组时,t 检验与 ANOVAs 在数学上是等同的。

数据是否符合 t 检验假设? 这包括各组方差同质性(即各组标准差大致相等)、观察的独立性、因变量(管理风格分数)在研究群体内的正态分布。如果考察了每组分数的频率分布,那么就会很容易知道是否符合同质性与正态性假设。也可以进行特定的统计检验来考察同质性与正态性(如考察数据分布的偏态或峰态)。如果数据符合这些假设,就可以开始分析数据。表 9.2 列出了最常用的参数统计检验的假设及是否违反假设条件。

举一个复杂的例子,考察领悟个体疗法是否比问题解决个体疗法产生更好的婚姻满意度。因变量为在婚姻满意度量表上的分数;两个自变量为治疗类型(两个水平:领悟疗法、问题解决疗法)和时间(三个水平:前、后、随后)。第一个自变量需要进行组比较;第二个自变量比较同一组被试在不同时间点上的差别。如图 9.1 所示,需要进行"混合 ANOVA",一个是被试间因素(治疗),一个是被试内因素(时间)。分析很可能是 2(治疗) ×3(时间)混合 ANOVA。

从专栏 9.2 可看出,虽然治疗组比较的结果很可能违反方差同质性,但时间因素没有违反方差同质性,因此需要考察方差同质性。可以从统计书中选一种检验方法来考察方差同质性(如 Keppel & Wickens, 2004; Maxwell & Delaney, 2004; 见本章后面的列表)。

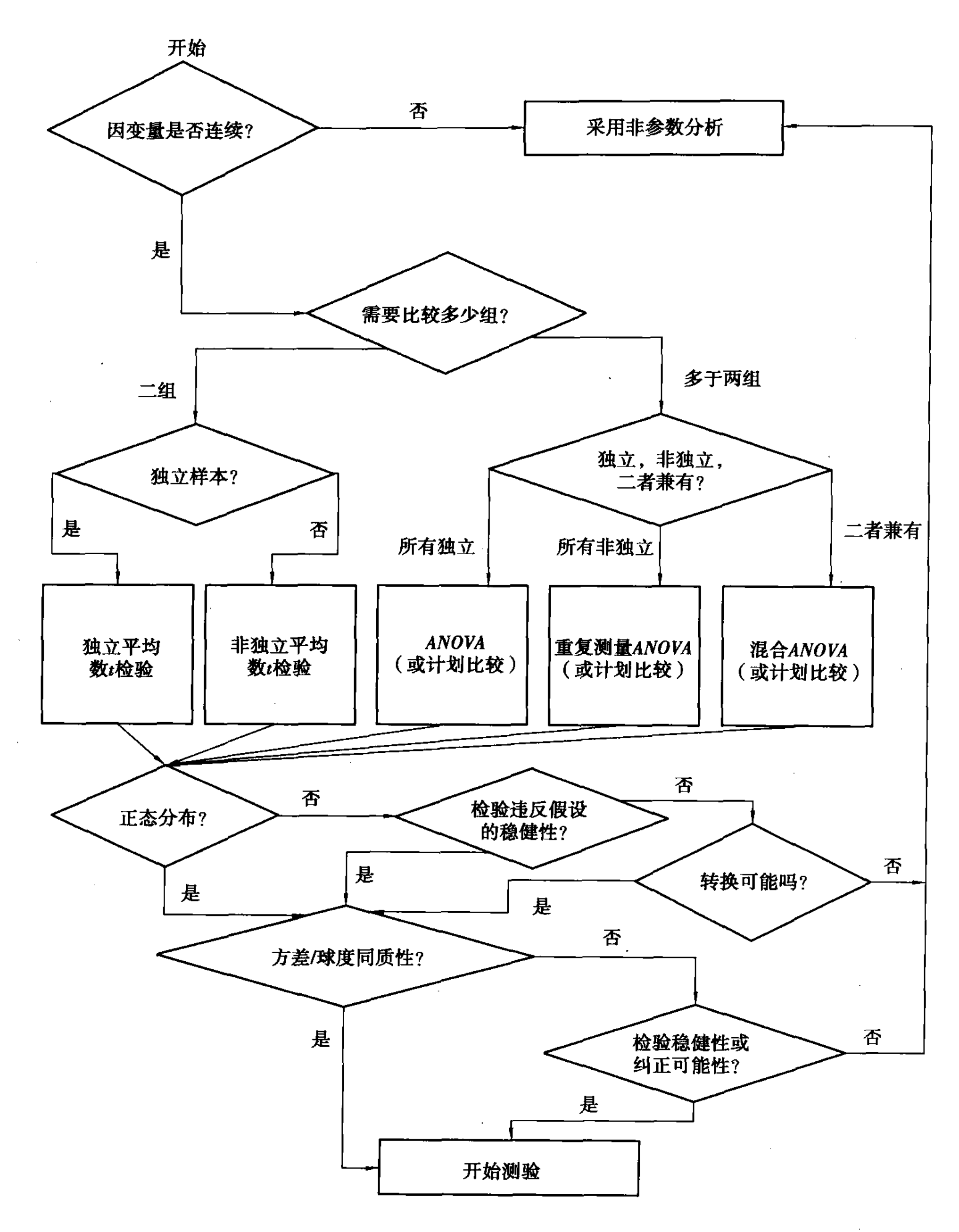


图 9.1　为组比较研究选择合适参数统计方法的流程图(假设分析单因变量)。ANOVA = 方差分析。

专栏 9.2

常见参数统计检验的假设

统计检验	假设
t 检验	方差同质性,潜在正态分布(通常都符合这两个假设,但重复测量不符合并且是不等的不显著[*ns*]),单元内的独立分数(不稳健)

专栏9.2 （续）	
方差分析	方差同质性，潜在正态分布；重复测量 ANOVA 也假设协方差和球形（圆形）同质性；ANOVA 通常是稳健的，但重复测量不是，并且是非常不等的 *ns*。小样本也可能不稳健（Tabachnick & Fidell，2001）；单元内分数的独立性（不稳健）
协方差分析	方差同质性，单元内分数独立性，潜在正态分布（对于相对相同的单元大小来说通常是稳健的，错误自由度 >19，双尾检验，无奇异值；Tabachnick & Fidell，2001），协变量与因变量间的线性关系，回归同质性（即每个单元内因变量与协变量的关系是相同的），能够可靠地测量到协变量（违反这个假设会导致检验力减少），强烈建议协变量与自变量的独立性；重复测量 ANCOVA 也假设协方差同质性
多元方差分析	多元正态分布，每个因变量与方差-协方差矩阵的方差同质性（如果单元大小不同就不会稳健），所有因变量间的线性关系，多重共线性（各因变量间）和奇异性（各因变量间）的缺失
皮尔逊相关	连续因变量；独立观察
多元线性回归	缺少多元共线性（不稳健）和奇异性，多元正态分布，方差齐性，预测值与效标之间的关系（可包含预测值间的交互作用），正确的测量预测值（违反此项会导致对回归系数[β]的低估），在分类上没有错误
结构方程模型	缺少多元共线性和奇异性，多元正态分布（获得稳健的卡方检验），线性关系（可以塑造其他类型的关系）

标准统计软件如 SAS 和 SPSS 能够进行上述多数检验。如果组方差非齐性，可以采用几种程序调整自由度来解决这个问题（见 Maxwell & Delaney，2004）。避免方差异质性带来的偏向会使得统计方法更保守（即减少犯 I 类错误的可能性）。也就是说如果不违反方差同质性这个假设，那么组间差异就会更大，更容易达到相同的统计显著性水平。另外，也可以选择 F^* 和 W 检验（Maxwell & Delaney，2004）。无论是对相等还是不等数量的单元，当方差不等时 F 检验的变式都能够产生没有偏向的结果。然而这并不适合包含重复测量的设计。

计划比较

流程图中的方框表明也可以选择计划比较，可对平均数进行配对比较，在收集数据之前进行。以此作为除标准 ANOVA 方法之外的另一选择。如果希望某些平均数在某些方向上不同并且只想检验这些差异，这时就可以用计划比较来代替小型的或整体的 ANOVA。

计划比较的优点是由于事先限制了检验的次数，因此统计检验力会更大一些，不足之处就是这很难随后考察不包含先验假设的数据。因为没有免费的午餐！确实，计划比较使统计检验力更强，代价就是随后的考察需要更严格的检验。总体来说，如果假设定义明确并且有研究支持，那么计划比较就是有意义的。当采用复杂

的设计,只需要比较很少的平均数并且不需要进行其他比较时(例如采用很多控制条件),也可以考虑计划比较。如果研究是探索性的,那么就没有必要为了更大的检验力而限制统计方法。多数探讨 ANOVA 的统计书也会探讨如何进行计划比较。需要注意的是计划比较有多种方式,取决于你如何选择比较组。

事后检验

如果不能做计划比较该怎么办呢?先别把统计书收起来!这时需要选择事后检验,它能够考察哪些平均数差异导致了显著效应。例如,之前说过的治疗结果研究,时间主效应显著,时间和治疗方式交互作用显著。需要进行事后检验来考察:①什么导致了时间主效应(时间主效应显著说明前治疗、后治疗和随后治疗之间存在差别,但没有说明是哪两个存在差别);②组平均数的哪些差别导致了交互作用效应。

交互作用意味着在自变量的各个水平上组间差别是不同的。在这个例子中,时间和治疗的交互作用可能意味着前治疗与后治疗平均数之间的变化在两种治疗方式中是不同的,或者后治疗与随后治疗平均数之间的差别在两种治疗方式中是不同的。6 种平均数(前-领悟,前-问题解决,后-领悟,后-问题解决,随后-领悟,随后-问题解决)中哪些存在显著差别,事后检验能够给出答案。

有很多种事后检验,可以从概念角度根据它们的检验力、违反假设的稳健性及每种方法最适合的情境来进行评估。马克斯韦尔和德莱尼(Maxwell & Delaney, 2004)很好地概述并比较了各种事后检验。被试内设计或不等-*n* 设计(unequal-n designs)没有稳健地违反方差齐性和正态分布假设,可以考虑适合这两种设计的事后检验。

ANCOVA 及 MANOVA

除了 *t* 检验和 ANOVA 统计之外,我们还可以考虑使用协方差分析(ANCOVA),多元方差分析(MANCVA)和多元协方差分析(MANCOVA)。接下来我们将分别讨论。

从本质上说,ANCOVA 是一种考虑到从因变量分离出来的一个或多个变量(称作协变量)的统计影响的方差分析。理论上,ANCOVA 体现的是在协变量水平上进行的 ANOVA,并且这些协变量是通过随机化方法得到控制的(实际上,这种情况更复杂)。当这种随机化控制不可行时,可用 ANCOVA 代替 ANOVA。这特别适合当因变量与另一变量有关,并且这个变量也与自变量有关的情况。记住,研究要考察的是某一特定自变量对因变量的影响,不想让其他一些不可控的变量影响结果。一个较好的例子就是种族与学术成就之间的关系。社会经济地位除了与种族存在关系外还与学术成就强烈相关,如果不控制社会经济地位,即使种族与学术成就之间不相关,也会得出二者相关的结论。

由于 ANCOVA 能够通过减掉由于协变量与因变量相关而产生的方差来减少误差方差(Tabachnick & Fidell,2001),这样就使自变量更容易产生显著效应。例如,考察 3 种不同的阅读计划对阅读分数的影响。然而 IQ 与阅读分数高度相关,它们之间的关系非常紧密,以至于如果不采取某种方法控制 IQ,就看不到 3 种干预之间

的任何差别。实质上，ANCOVA 把 IQ 作为协变量能够把 IQ 与阅读之间的关系从分析中提取出来，使得更容易检测到自变量效应。

当变量间存在潜在混淆（如种族与社会经济地位），想从统计上分离出其影响时可采用 ANCOVA。想想之前的例子，比较领悟指向疗法与问题解决疗法的治疗结果的研究。设想即使随机分配了被试，但领悟指向疗法组的教育水平仍显著高于另一组，并且教育水平与治疗效果相关；在这种情况下，领悟疗法组的治疗效果好可能是因为教育水平而不是因为治疗。可以通过把教育水平设为协变量来避免这个问题。理论上这能够移除教育水平对婚姻满意度的潜在影响，进而更好地考察治疗效果。然而在这个程序的适当性上存在一些争论，因为 ANCOVA 在某些情况下可能会产生很难解释甚至误导性的结果，这将会在后面讨论。

采用 ANCOVA 时应该注意一些问题。首先，只有当协变量与因变量存在显著相关时，采用 ANCOVA 来控制潜在混淆才是有意义的。如果不存在相关，那么可能存在的混淆变量不止一个。这就不能够为治疗组之间的差别提供可行的解释。此外，ANCOVA 与 ANOVA 结果不会存在太大差别（但会失去一个自由度，即引入协变量能够增加一点检验力）。其次，ANCOVA 只能控制协变量与因变量之间的线性关系，如果协变量与因变量间存在曲线关系，混淆仍存在。第三，当协变量与自变量间的差别存在相关或协变量由自变量的差别引起，采用 ANCOVA 会使结果解释变得困难（见 Maxwell & Delaney，2004）。第四，如果没有随机分组（即不是真实验设计）而采用 ANCOVA 也会导致解释困难，胡伊特马（Huitema，1980）对这个问题进行了详细探讨。或许由于这些解释上的问题，一些人认为 ANCOVA 只能够用于提取因变量与协变量间的方差，并不能够控制非预期的（或预期的）混淆。因此，虽然 ANCOVA 是一个强大且有用的程序，但并不能直接解释结果。如果某个研究采用 ANCOVA 进行分析，那么需要考察它能做什么及不能做什么。Huitema（1980），Maxwell 和 Delaney（2004），及 Shadish 等（2002，第 8 章）都涉及了 ANCOVA。此外，Huitema（1980）提供了一些能够用于 ANCOVA 的事后检验公式。

除能够同时分析多个因变量外，MANOVA 与 ANOVA 是类似的。例如，采用 MANOVA 考察领悟指向与问题解决研究，因变量有交流满意度、性互动和帮助任务。MANOVA 会问"各种因素（自变量）在因变量**区组**（group）内会存在差别吗"。在概念上，MANOVA 程序在分析中创造了一个联合所有因变量信息的合成变量，然后分析这个合成变量，最后报告这个合成变量的主效应和交互作用是否显著。结果能够说明自变量是否影响这个合成变量，但不知道它是否影响组成这个合成变量的单个变量。

研究者们常常通过 MANOVA 来减小 I 类错误的比率。当对许多独立的因变量进行统计分析时 I 类错误会增加，进而会增加仅仅是因为检验次数的增多而导致显著的可能性。由于 MANOVA 把多个因变量看成一个组，减少了分析次数，进而减少了 I 类错误的可能性。如果 MANOVA 显著（符合逻辑），研究者可以继续在每个因变量上进行一系列单个的 ANOVA，看看哪个是显著的，这类似于 ANOVA 显著后再进行事后检验来比较特定组。另外，也可以逐个排除变量来看看是哪些变量导致了显著效应（Tabachnick & Fidell，2001）。

休伯蒂和莫里斯（Huberty & Morris，1989）对这种方法的逻辑提出置疑，认为

MANOVA 并不能始终避免 I 类错误,并且 MANOVA 主要适用于研究者关注因变量作为一个系统,或因变量在概念上相关,并且想评估相同或相似的潜在构念的情况。多个 ANOVA 适用于因变量在概念上独立、探索性研究、前人研究采用独立的单变量分析(这样可以和前人研究进行对比)和考察组之间是否等值时(如,可能存在混淆变量)。我们的建议是,如果因变量在概念上相关,当符合假设或某一特定类型的 MANOVA(如重复测量)稳健地违反假设时,MANOVA 才可能有意义。如果因变量在概念上不相关,可以通过纠正 α 水平来避免 I 类错误。如果在这种情况下仍选择 MANOVA,要确保有很好的理由并且能够反驳休伯蒂和莫里斯(Huberty & Morris,1989)的观点。

非参数统计

如果数据不符合 ANOVA 假设,这该怎么办呢? 可采用非参数统计。如果打算进行被试间 ANOVA,但是你的序数数据不适合,则可以选择克鲁斯凯-沃利斯检验(Kruskal-Wallis test)这种非参数检验,它适合序数数据,当只比较两组时与威尔克森·瑞克·桑姆和及曼恩-惠特尼 U 检验等值(Maxwell & Delaney,2004)。虽然克鲁斯凯-沃利斯检验没有有关潜在分布的假设,但它假设方差齐性。因此,马克斯韦尔和德莱尼(Maxwell & Delaney,2004)建议当方差非齐性时采用 F^* 和 W 检验,特别是当单元大小不等时。联合等级或配对等级检验(Maxwell & Delaney,2004)可对单元进行配对比较,就像在 ANOVA 中进行事后平均数比较一样。

如果采用重复测量并且只有一个(重复)自变量,可采用傅莱德曼检验(Maxwell & Delaney,2004)。和克鲁斯凯-沃利斯检验一样,它分析的是被试分数的等级顺序。对于分类数据,最合适的分析方法是卡方检验或交叉分类分析(Rodgers,1995)。当自变量有多个水平或有多个自变量时,若卡方检验显著,可进行与事后检验具有同样效果的检验,如 z 验,或者采用与 ANOVA 策略具有同样效果的更复杂的非参数检验(模型分析或对数线性分析)。

相关统计分析

如果想考察变量在一组被试内的关系,而不是比较它们在不同组中的差别,那么第一步是选择相关统计方法。如果有两个变量,可采用二元相关。如果考察变量集(即多于一个变量)与一个或多个不同变量间的关系,回归或模型检验是最好的选择。

二元相关(参数与非参数)

测量相关的最简单方法是二元相关系数。如果在两个变量上的数据是等距的或是等比的,可采用皮尔逊积差相关(r)。记住,皮尔逊相关适用于等级排序,不能够系统考察两组数据间的平均数差异。例如,考察两个个体各自观察同一个人晃头频率所得数据的相关。不仅关注他们是否认为那个人的晃头频率更高或是更低,也应关心他们同时观察时是否报告了相同的晃头频率。在这种情况下,可能需要考察

平均数差别的相关(如,组内相关系数;von Eye & Mun,2005)。

如果一个变量是序数数据,另一变量是序数、等距或等比数据,那么斯皮尔曼等级相关比皮尔逊相关更适合。如果其中一个是二分变量,可采用点二列相关。如果两个都是二分变量,那么二者之间的相关为 *phi*(Φ)系数。注意用于计算 *r*、点二列和 *phi* 系数(作为特殊情况)的公式,让计算机做二分数据的 *r* 值会产生一个纠正的点二列或 *phi* 系数(Cohen, Cohen, West & Aiken, 2003)。

回归策略

回归分析中有一个**效标变量**(criterion variable)(有时也叫**因变量**[dependent variable]或**结果变量**[outcome variable])及一个或多个**预测变量**(predictor variables)(或**自变量**[independent variables])。有两个或多个预测变量时,回归程序会产生一个公式用于描述联结各个预测变量的最好方式。这会告诉我们在预测变量上的分数能够预测在效标变量上的分数的程度并用公式把它们联结起来。可采用回归分析来考察对治疗师某一特定行为的等级评定与治疗结果之间的关系;年龄、社会能力、研究生考试(GRE)分数及本科平均绩点(GPA)是否与研究生平均绩点或专业执照考试分数相关;各种人口学变量和认知测量结果是否与从癌症治疗中康复的几率相关;等等。虽然回归可以只有一个预测变量,但常常包含多个预测变量。因此,回归的参数形式也叫做**多元回归**(multiple regression)。

参数多元回归

如果对被试在预测子上的分数一无所知,多元回归分析的参数形式会产生一系列能够提高预测能力的变量。这种分析程序会考虑到一些预测子间的相关,因此是多余的。每一个显著的预测变量会提供独特的(不是多余的)信息(即与前一个预测子提供的信息不同)。可用二分变量和连续变量作为预测子,但在多元回归的参数形式中,效标必须是连续变量。(稍后会提到回归的非参数形式[对数分析和逻辑回归])具有多于两个水平的分类变量也可作为预测子,但需要特别处理(见 Cohen et al., 2003)。更复杂的回归变式能够评估不同预测变量之间的相互作用(即一个变量是否能够调节另一个变量与效标变量之间的关系)。艾肯和韦斯特(Aiken & West, 1996)描述了如何生成、分析和解释回归分析中的交互作用。

有多种不同形式的回归,多数与预测变量输入方程的顺序有关,逐个把预测变量输入到一个或多个组中。不同形式的回归在检测变量能否在预测效标变量时提供新信息(不同于方程中其他变量包含的信息)的方式上也有差别。

在**层次多元回归**(hierarchical multiple regression)中,根据理论和方法学知识,事先指定回归方程中变量的顺序。分析首先提取出第一个变量的方差量,然后提出第二个变量的方差量,以此类推。可以指定每个变量的输入顺序或组变量输入子集的顺序然后输入这些子集,并整理出子集在每个集内的顺序。

例如,我们的一个学生想根据父母的婚姻满意度及父母与孩子的关系质量(凝聚力)来预测姻亲满意度。“姻亲满意度”是效标变量。他首先把几个人口学变量(作为一个集)输入到层次多元回归分析中,然后输入一个社会期望测量分数。最后,输入婚姻满意度和凝聚力测量分数(作为一个集)。他选择这个的顺序是为了

控制人口学特征及在自我报告测量中以社会期望的方式进行回答的倾向。同时也想保持人口学和社会期望变量恒定,来考察婚姻满意度、凝聚力和姻亲满意度之间的关系。换句话说,考察除人口学与社会期望测量外,婚姻满意度与凝聚力分数能否提供其他新的信息。不过没有必要对集内变量进行排序。

层次多元回归有时也可用于检验调节假说,即两个或多个变量间的关系可用第三个变量来解释,这第三个变量称为"调解者"。例如,设想低学术成就与高辍学相关,并且这种关系受到学生对学校的不良态度、与老师缺少沟通及缺少学校活动这3个因素的调节。巴伦和肯尼(Baron & Kenny,1986)和霍姆贝克(Holmbeck,1997)描述了采用回归分析来检验这种调解假设的经典分析方法。

麦金农,洛克伍德,霍夫曼,韦斯特和希茨(MacKinnon,Lockwood,Hoffman,West & Sheets,2002)认为这种方法缺少检验力,然而他们提倡的其他方法也是利用了回归策略。

除层次多元回归外的另一方法是**逐步回归**(stepwise regression)。在这个程序中,纯粹地计算决策规则,识别出哪个变量能够最好地预测效标变量,除此之外哪个变量能够显著预测效标变量,等等。在层次多元回归中需要人为地事先安排顺序,而在逐步回归中由计算机进行排序。如果在之前的例子中采用逐步回归程序,只需要输入所有变量并让计算机指出在预测姻亲满意度时预测变量的顺序。

排序有多种方式,需要告诉计算机采用哪个程序。包括**前输入**(forward entry)程序,**前逐步**(forward stepwise)程序,和**后删除**(backward deletion)程序。多数逐步方法采用的程序能够增加和删除变量,并且能够检验每个变量在分析时的表现。至关重要的一点是回归分析能够以多种方式进行。根据分析目的来考察选项及其利弊,然后作出明智的选择。

科恩等(Cohen et al. ,2003)指出了逐步回归程序的几个不足。第一,当预测子很多时,此种方法的预测估计在随机水平。第二,采用此程序得出的结果不适用于其他样本。第三,由于输入和删除预测变量的规则不同,有些结果可能具有误导性。第四,采用逐步程序意味着可以不用考虑预测子与效标变量之间的逻辑关系。这有点像采用ANOVA和事后比较来分析数据而不必事先计划和指定考查内容。

逐步程序适用于:①大样本;②可重复(另外采用一个样本或把样本随机分成两组);③研究目标不是基于理论(在一些应用性研究中,研究目的可能只是简单地预测某一现象而不是进行解释,Cohen et al. ,2003)。塔巴奇尼克和菲德尔(Tabachnick & Fidell,2001)认为逐步程序是模型建构技术,在某一领域的早期研究中会采用这些方法。他们把层次程序看作是假设-检验技术;①当一个研究领域足够成熟,可以进行合理预测时可采用这种方法。

与所有参数检验一样,回归程序如果不满足某些条件就会产生问题。第一,多元回归程序假设不存在多元共线性。当两个或多个预测变量间高度相关时存在多元共线性,在这种条件下会产生不稳定的回归方程(即每个预测变量的权重不会在

---

① 可以联合逐步程序与层次程序。即,以层次的方式输入一系列变量子集,但采用逐步程序来指定每个子集中变量的顺序。

另一个新群体中得到重复)。①如果存在多元共线性问题,可以从相关变量中选择最重要的预测子或者看看能否把相关的变量合理结合成一个分数。

第二,预测变量不能够与其他预测变量组合(这被称作**奇异性**)。换句话说,不能把冲撞分数、谩骂分数、拳击分数和攻击总分作为预测子,最后一个是前三者的组合。第三,回归假设每个预测子与效标之间存在线性关系(但多元回归也可用于检验预测变量间的交互作用)。第四,多元回归假设潜在的多元正态分布。这意味着不止一个变量正态分布,所有可能的预测子组合都是正态分布。第五,假设方差齐性,回归方程产生的预测在回归线上的不同点是类似的。可以通过考察残差来检查后3个假设,残差分析能够看出回归方程的每个预测分数与实际分数之间的差别(每个差别都被称作**残差**[residual])。回归线上不同点存在不同类型的残差能够影响多元回归的假设(见 Tabachnick & Fidell,2001)。

多元回归的最后一个假设很难得到满足。首先,回归假设包含了所有相关预测子并且没有预测子是与效标不相关的(Licht,1995),达不到这一条被称作规格错误。另外,多元回归程序假设对变量的测量是准确无误的。

这些假设看起来不可能得到满足:如果知道哪些变量与效标变量相关,哪些不相关,还需要做研究吗?并且有多少变量可以得到准确无误的测量?事实上,虽然没有人可以做到完全符合这些假设,但研究者们仍会采用回归分析。切合实际的建议是考虑符合这些假设的程度而不是忽略它们(Klem,1995)。利希特(Licht,1995)建议采用理论和前人数据来选择合适的预测变量,采用最可靠且有效的测量方法来把测量错误降到最低。

回归的非参数形式

逻辑回归(logistic regression)与多元回归类似,即多个预测变量与一个效标变量相关。但与多元回归不同的是,预测子可以是连续变量也可以是非连续变量,效标变量是分类变量,而非连续变量。另外,逻辑回归的数学分析更符合非参数统计(与参数统计相比),并且与多元回归的参数形式有很大不同。因此,非参数回归统计考察的是某一特定单元中的频数与预期频数的吻合程度,而不是每个预测变量能够预测效标的权重。对数分析(对数-线性分析的变式)比逻辑回归的限制更多,因为所有的变量(预测变量与效标变量)必须是分类变量。

差别函数分析

差别函数分析(DFA)与逻辑回归有许多相同的功用,但 DFA 属于参数统计,最适用于从变量集中(自变量或预测子)预测离散的组成员(因变量)。例如,评估几个 MMPI 分数能够预测精神病诊断的程度,或联合人口学数据与评估工作和公司态度的分数来预测未来员工在工作的第一年是否会辞职。

这些问题类似于组比较问题,因为需要把样本分成两个或多个离散组来形成因

---

① 一个共同准则是高于0.7的二元相关意味着多元共线性问题(例如,Tabachnick & Fidell,2001)。然而低于0.70的二元 r 并不说明不存在多元共线性:单独看时,两个预测变量可能与第三个预测子只存在中等程度的相关,但组合起来时就可能存在高相关。SPSS 和 SAS 能够通过计算指标(如容许水平[tolerance leves])来检验变量组合的多元共线性;科恩等(Cohen et al.,2003)也讨论了这个问题。

变量(如,反社会人格障碍组与控制组;离开公司的员工与留下的员工)。它们类似于回归问题,因为需要考察变量组合能否把样本分成各个组。如果 DFA 符合假设(见专栏 9.2),它在这些情境下会很有用;否则,逻辑回归可能是更可行的选择。

实质上,DFA 考察预测子(自变量)并制定方程来衡量每个变量,以便更准确地把被试分组(因变量)。与回归一样,DFA 存在变异。DFA 与 MANOVA 在数学上紧密相关。需要重点考虑是否需要计算机优先检验某些变量(层次 DFA)或者是否需要计算机对变量在统计上进行排序(逐步 DFA 或直接“标准”DFA,见 Tabachnick & Fidell,2001)。

因素分析

我们之前假设一些变量是自变量,一些是因变量,并且关注自变量与因变量的关联程度。然而还可以不根据自变量和因变量来考察数据。如果关注那些没有指定因果或预测关系的变量,可以采用其他相关统计方法。因素分析及其变式,主成分分析这三者都是变量集内相关。研究者们常常采用这些方法来把众多的变量或项目集减少到更小数目或检验变量间的潜在结构。例如,一个研究者开发了一个新问卷并想在考察组差别之前把 50 个项目减少到几个同质子量表中。

使用因素分析时需要考虑大量且复杂的问题(见 Fabrigar, Wegener, MacCallum & Strahan,1999; Floyd & Widamon,1995;Preacher & MacCallum,2003 对一些问题的精彩回顾)。然而一定要说的问题是:样本大小。普遍的经验法则是每个项目(变量)至少需要 5 个被试或一共需要 100 个被试,甚至更多。麦卡勒姆,维达曼,张,洪(MacCallum, Widamon, Zhang & Hong, 1999)对此提出了置疑,认为被试数量取决于因素结构能够解释项目方差的程度,或能够通过合理数量的项目测量到每个因素的程度。如果这个分析能够解释合理数量的方差(即公因子方差约为 0.50)并且每个因素至少有 6 或 7 个项目;根据麦卡勒姆等的观点 100 ~ 200 个样本数量足够了。每个因素中的项目数量越少,需要的被试数目就越多。另外,许多研究者建议为了保证结果的可靠性最好重复一次因素分析。你愿意招募很多被试来进行正确的分析吗? 如果不愿意,就不要认为因素分析能够解决多变量问题。一个更好的解决方法是首先考虑是否所有的测量都是必要的,特别是首先要确定它们测量的是不同的构念。

另一个需要考虑的问题是采用探索性因素分析(CFA)还是验证性因素分析。探索性因素分析由计算机决定变量间的最佳组织方式,这在概念上可能有意义也可能没有意义。主成分和主轴因素分析都属于探索性因素分析。相反,验证性因素分析是一种模型检验技术,指定哪些变量负载哪些因素,然后根据实际数据来检验理论模型。研究者一致认为探索性因素分析最适合开发测量工具的初级阶段,而验证性分析最适合采用成熟的测量工具来检验明确的理论中项目或分数是如何组织的(Floyd & Widaman,1995)。

模型-检验方法

模型-检验方法与本章前面介绍的其他统计方法不同,它检验的是实际数据是否与研究者提出的理论模型相符。心理学中最著名的模型-检验分析可能是结构方

程模型(SEM),CFA 是 SEM 的一个应用。采用 CFA 时,需要告诉计算机问卷中的不同的分量表由哪些项目组成,模型检验分析会指出这些分量表是否由这些项目组成。

SEM 也可作为多元回归的另一选择来检验调解假设。举个例子,态度和心情调节低学术成就与高中辍学之间的关系。

假设这个例子中存在相当复杂的关系集——即成就与辍学相关,成熟也与不良态度和与他人缺少联系相关,并且态度和心情也分别与辍学相关。另外,如果知道有关学生态度和联系的信息就能够解释成熟与辍学之间的关系。SEM 能够在一次分析中检验所有的关系。

SEM 也可用于检验更复杂的模型。例如,SEM 逐渐取代了路径分析。路径分析考察预测子与效标变量之间的直接和间接关系。典型的简单路径分析推测 A 导致 B,B 导致 C,A 也直接导致 C。此模型认为 B 能够解释 A 与 C 之间的间接关系,但 B 不能够解释 A 与 C 间的直接关系。调节分析是路径分析策略的子集。也有复杂的路径模型。例如,A 导致 B,B 导致 C,C 导致 D。这种类型的模型通常采用回归方法来计算直接和间接路径系数,路径系数体现的是变量间的关系(Klem,1995)。若样本足够大,现代研究中常采用 SEM 来检验这种类型的模型。

当采用多种测量工具来测量同一构念,并且想考察构念间的关系(有时被称为“潜变量”)而不是估计实际测量变量(也被称为“观察变量”或“指标”)时,研究者们也采用 SEM。这是因为,从概念上说,SEM 包含了因素分析(根据观察分数来创建合成变量或潜变量)与回归分析(考察一个变量或变量集能够预测另一变量的程度)的特征。采用这种类型的模型时需要告诉计算机:①哪个变量测量哪个潜构念;②希望构念间如何相关。模型与数据间的拟合程度取决于①与②。

模型-检验方法也需要事先告诉计算机你希望变量间如何相关。采用分析策略来检验预测的关系与数据中存在的共变形式拟合的程度。可采用卡方检验,但在这里显著的结果意味着预测显著偏离实际数据。实际上,这是希望从统计中得到非显著结果的罕见情况之一。

然而这种方法存在的问题是卡方是否显著取决于模型与实际数据之间的差别及样本大小。样本越大,卡方检验到的差别越小。然而统计学家们通常建议 SEM 采用大样本。这会增加模型不符合实际数据的机会,虽然缺少拟合度的机会很小。只关注卡方检验的结果会导致过早地拒绝预期的模型。

幸好研究者们与统计学家们意识到了这个难题,并设计一些解决有关大样本/显著卡方问题的方法。解决办法是使用一些好的拟合指标(见 Ullman,2001,描述了不同类别的拟合指标)。当然,这意味着需要选择合适的指标。许多人建议要采用多个指标。每个指标都有其自己的分界分数(见 Raykov & Marcoulides,2000,第一章,常见指标的经验法则)。

结构方程模型有几个优点。最大的优点是可在一次分析中同时考察多个关系,而不必进行多次分析,这会减少 I 类错误。另一个优点是需要事先阐明想要检验的理论或模型。最后,如果对同一构念采用多个指标(一些人建议每个构念至少 3 个指标),用 SEM 检验构念分数与其他变量间的关系,这需要从模型中删除一些测量错误,以增加统计检验力。然而 SEM 并不是适合所有的数据集。首先也是最重要

的一点是样本问题。一般说来,模型越复杂,期望的效应越小,缺失数据越多,数据的偏态越严重,需要的样本越大。根据我们的经验,已发表且采用 SEM 的研究中,被试数量很少有少于 100 ~ 150 人的。厄尔曼(Ullman,2001)认为简约模型中 200 个被试"应该是合适的"(p. 659);汤普森(Thompson,2000)也建议每个观察变量至少需要 10 ~ 15 个被试。其次,初学者可能喜欢采用 SEM 来检验复杂模型。这不仅需要大样本,也容易产生各种各样的问题,如计算机输出结果为"不能识别此模型"及缺失关键统计,或在其他方面看起来有些奇怪。

我们强烈建议参加一些包括大量 SEM 的课程,获得一些使用 SEM 的经验,开始收集并阅读关于此方法的文章和书籍(如 Thompson,2000),如果是初学者并且在论文中可能会使用 SEM,最好列出有关 SEM 的参考文献。

设定 α 水平

如果已经选择了一个分析计划,那么就准备开始了,对吧?不尽然!看看分析计划需要多少步来检验假设。3 次还是 30 次?如果更接近后者,记住进行多次的统计检验会增加犯 I 类错误的可能性,这个问题叫做实验误差率。

有几种解决办法。第一,看看是否真的需要进行这么多次统计检验:是否有其他可行的方法能够运用较少次数的分析来检验假设?例如,可以采用 SEM 代替回归方程或采用回归代替相关。第二个办法是重新检验假设,看看在一个研究中是否有太多的假设。如果是这样,减少假设的个数,选择最重要且可行的假设(见 Wilkinson & The Task Force on Statistical Inference,1999,相关讨论)。第三种常见办法是根据检验次数来降低 p 值,可采用 Bonferroni 或其他方法(见 Keppel & Wickens,2004),但这会降低统计检验力。所以需要用新的 p 值来重新计算样本大小(检验力分析)。凯佩尔和威肯斯(Keppel & Wickens,2004)建议的另一个方法是仔细考虑假设。如果每个假设都独立且真实地测量某物,那么可把所有分析的 α 水平设为 0.05。相反,如果你的假设形成了族(例如有 5 个假设,其中两个在逻辑上相关,另外 3 个形成一个独立的"族"),那么需要根据族来调整 α 水平。

考察实验错误率需要哪些分析呢?多数学生会做 3 组分析:初步分析,假设-检验分析,及补充分析(进一步开发数据,用于解释假设-检验分析结果)。

考察 I 类错误时,主要关注假设检验。一般建议初步分析时把 α 水平设为 0.05,因为这些分析的目的在于寻找混淆或问题,此时Ⅱ类错误比 I 类错误更重要(当实际上存在几个潜在混淆时协变量能够控制这些混淆)。在这些情况下 p 值应该增加而非降低。补充分析时,可以调整 p 值也可以不调整。如果进行谨慎的探索性分析,把 α 水平降低;相反,如果害怕错过某些重要的东西,就把 α 水平设在 0.05。对于后者,效应小或者中等时不要过分解释孤立的结果——这很可能是假的,而应寻找相关补充分析中的一致性和大效应,这时产生随机结果的可能性更小一些。

可以看出,选择有多种,全面考虑你想采用的方法并把它加入到分析计划中。

小心非独立数据

虽然多数问题和数据类型都有合适的统计方法，但特定的环境会使统计复杂化。首先许多参数和非参数统计的一个基本假设是：观察的独立性。凯佩尔和威肯斯（Keppel & Wickens，2004）对此进行了明确定义："这个假设是指一个人的行为不会影响到另一个人的行为"（p. 134）。这个假设的重要性在于违反它会使结果出现偏差，导致低估统计标准误，会增加 I 类错误的可能性。

在被试内设计中，对来自同一被试的因变量的测量次数超过两次时就会违反此假设。一个常见例子是治疗组与控制组在前测中都被测量了某一研究者感兴趣的因变量，接受实验或控制操纵后，在后测中又对此因变量进行了测量。

进行组比较时（如对非独立样本的 t 检验，重复测量 ANOVAs）比较完善的方法都能够处理这个问题。回归分析可从两个时间点处理纵向数据，把时间 2 上的数据作为因变量，时间 1 上的数据作为预测变量或控制变量。更复杂的模型检验方法，如层次线性模型（Raudenbush & Bryk，2002；Singer & Willett，2003），也可以处理在多个时间点收集的纵向数据。然而在 3 种情境下处理非独立数据更复杂。首先，重复测量非参数分析不常见。例如，考察在不同的实验情境下同一被试群体会选择 3 种产品中的哪个。因变量是分类变量（被试选择哪种产品），但数据不是独立的，因为每个被试都接触到了 3 种产品。虽然在这种情境下需要采用一些统计方法，但我们不知道应该采用的是哪种，只能说在我们仔细读过的教科书中没有说明。

下面是另一个复杂的情境。被试不止接触到一个条件（创建一个重复测量设计，一个自变量是一个处理条件，因变量是一个重复测量变量），并看看在另一个连续变量上存在差别的被试在对操作的反应上是否有差别。例如，我们的一个学生，劳拉·霍耶（Laura Goyer）让男孩看两组录像：一组讨论"正常的"日常话题，另一组讨论异常活动。观察男孩在看两个不同组讨论时的情感表达。她关注的是行为不良的男孩是否对这两种录像的反应不同。在理想情况下，她想考察行为不良（连续变量）与录像类型（被试内实验变量）的交互作用。由于不能确定一个合适的统计方法，劳拉采用行为不良测量工具，按行为不良水平把男孩分成高、中、低 3 组。然后分析就简单多了：进行 3（行为不良水平）×2（讨论类型）混合 ANOVAs，其中讨论类型是一个重复测量变量。

第三种情况是研究者往往意识不到他们违反了独立性假设。回忆一下凯佩尔和威肯斯（Keppel & Wickens，2004）对独立性的定义：一个人在因变量上的表现不影响另一个的表现。在一些被试间设计中，虽然每个被试只接受一种实验处理，但这并不能保证独立性。

当被试来自同一群体，并且不可避免地受到其他人影响时就会违反独立性原则。例如，研究者们观察约会情侣或小群体内个体间的相互作用。一个人的行为明显会影响另一个人的表现，不能认为从情侣那儿获得的数据或二者间的交互作用是独立的。当班级或集体作为一个群体接受干预时，一个人对治疗的反应很可能会影响到其他人的反应。几个家庭成员参加研究，组内个体相互作用，以及对群体进行

干预,在这些情况下数据不可能是独立的。最常见的解决办法是进行多水平或多层次分析(如 Raudenbush & Bryk,2002)来处理非独立性问题。然而这种分析方法很复杂,可能需要专业软件。

对非独立数据的分析方法不像独立数据那么多、那么成熟和方便。事实上所有的统计学家都建议必须认真考虑独立性假设。因此,最开始就要确保能够处理非独立性问题,这可能需要重新设计研究。另一个方法是把被试内独立变量转化成被试间变量。另外,可以避开独立性问题来分析数据。例如,可以对情侣的数据进行平均,把"对"作为分析单元。最后,可能需要寻找并学习不熟悉的方法。

提防因果术语

不要根据统计方法的名称或性质来推测因果性。例如,一些学生错误地认为显著的 ANOVAs 意味着自变量"导致"因变量的变化,但并不是这样的。确实,自变量的不同水平与因变量上的差别有关。然而,自变量是否导致因变量上的变化取决于研究设计,而不是用于检验变量间关系的统计方法。如果不操纵自变量,也就不存在因果关系。

在这种情况下,一些人把结构方程模型称作"因果模型",这个称谓产生了大量不恰当的推论。多数 SEM 分析的数据来自一个时间点。同样,基于横向数据的路径分析只是一种相关分析。回归分析是指一个或多个变量"预测"一个结果并且所有的变量是同时测量的,调节分析同上。无论用于描述统计的行话是什么,相关只是相关。研究设计,而非统计方法,决定着自变量是否与因变量存在因果关系。

谨慎征求意见

此时,你可能知道哪些统计方法适合自己的研究,可能也想知道自己选择的统计方法是否正确,可以去图书馆或好的网站详细阅读这些统计方法,可以请教其他同学或老师的看法,看看你们的选择是否一致。

如果仍感到不知所措且需要额外的帮助该怎么办呢? 你是否会为了某些目的而寻找帮助或征求意见? 对这些问题的回答取决于征求意见的程度。"征求意见"包括问导师、同学或学院的统计学教授有关分析中的几个特定问题,进而考虑是否雇用某个人来为你选择、编程并进行分析。顾问和自己(想要获得学位的人)所做的工作量在这几种情况下是不同的。

征求意见是有益的。多数教师会咨询在某一特定统计程序方面比自己懂得更多的同事,并从他们那里获得经验。经验丰富的研究者常常把统计咨询写人预算中。在写这一章时我们也咨询了几个同事。向教师和研究生咨询某一特定的统计分析方法能够为你指引正确的方向,通常是大家所能接受的。事实上,找出自己不懂的东西,然后通过阅读和请教同事来找出答案,这是专业研究的一个重要部分。多数学生会在选择统计方法或证实他们的选择时寻求帮助。

除了咨询他人还有什么呢？这是个复杂的问题。在理想状态下，学生们自己做了所有的工作，只是偶尔征求他人意见。不幸的是，一些学校中的统计学课程并没有把学生教好。另外，学生在使用器材时存在大量困难，导师可能与学生一样恐惧统计，或者学生得不到用于复杂分析的计算机设备和参考资源。在这些情况下，让他们在没有帮助的情况下弄明白如何处理和分析数据是不现实的。

无论使用何种咨询手段，记住最终是自己要对研究中的统计负责。这意味着你一定要知道为什么一种统计方法合适而另一种不合适，在分析中应该采用哪种计算机程序并进行哪些选择，如何处理缺失数据，计算机的输出结果是什么意思，分析是否正确，及如何解释结果。换句话说，你是否真正实施了分析，必须知道自己都做了什么，从内到外，并对结果的准确性及其解释负责。

如果需要更多的咨询该怎么办呢？首先，我们建议尽可能地独自分析。如果有合适的计算机资源但是不理解某一特定统计方法，可以咨询他人有关如何选择统计方法、编程和解释结果。可以自己记分、删除数据、输入数据并进行分析（或许电脑顾问能够帮助你找出程序中的错误）。其次，我们建议你寻求有关技能问题的辅导。此时用**辅导**（tutoring）这个术语比**咨询**（consultation）更适合。在辅导快结束时，应该检验一下自己做了什么，为什么这么做，所采用的统计方法的假设和特征，及自己的发现。

事实上，这种检验应该出现在答辩中。“我真的不确定做了什么，我得问问做过这种统计的人”，这不是在最后答辩中回答统计学问题的方式。你，并且是你自己一个人——而不是帮助过你的那个人——必须在委员们面前进行论文答辩。委员会评估的是你做研究的能力，而不是你的顾问。

是研究者自己，而不是提供帮助的人要获得学位。因此，在获得他人帮助时要明白自己在做什么，尽可能自己做并对统计方法的选择和计算准确性负全部责任。

即使有人帮忙做了所有的统计分析，也要自己准备数据。在顾问告诉你如何做之后，自己把数据输入到计算机中。这样可以知道原始数据是什么样的，也会促使你去想一些重要问题，如怎样处理缺失值。你也可能会检测出奇异值——可能会曲解结果的极端值，这需要谨慎的处理。我们也建议你自己做初步的分析，看看平均值、标准差、频数分布等。如前所述，频数分布有助于确定哪种统计方法合适，提醒你改变分析计划。我们会在第10章详细说明初步分析。

若仍然困惑，寻找额外的帮助

阅读了本章之后仍然不知道选择哪种统计方法？可能不止你一个人会这样。每个人对统计材料的理解程度不同。书籍、教授及专业学科在谈论分析时会采用不同的语言系统。不仅符号不同，有时人们对同一分析会采用不同的名字。本章不能包含每个统计方法或基本方法的每一个变式。

毫无疑问，此时许多人需要额外的帮助。制定决策时可以采用决策树，就像专栏9.1那样，基于统计课本中的信息汇编和其他决策辅助，包括本讨论中所涉及的内容。耶茨（Yates，1982）采用表格的形式提供了类似的信息来帮助同学们选择合

适的统计方法。安德鲁斯,克莱姆,戴维森,奥马利和罗杰斯(Andrews, Klem, Davidson, O'Malley & Rodgers, 1981)提供了一个内容丰富的决策树用来选择合适的统计方法,包括相关、非参数和参数检验,为每个统计方法提供了参考,并提供了相应的统计软件包。然而,他们只简单提及了重复测量(被试内)统计方法。塔巴奇尼克和菲德尔(Tabachnick & Fidell, 2001, pp. 27-29)提供了一个多元技术决策树,包括相关统计策略。

另一种办法是考察影响因子较高的杂志上最近与自己研究相近的文章,看看他们都采用了哪些方法。找那些与自己的研究在变量和测量类型(例如,适合参数分析或者反之)上相似的例子,他们也应该与你具有相似的研究设计。

我们建议使用工具时就像阅读本章一样,把它作为获得想法的一种途径,可以随后通过阅读和咨询验证或否定这个想法。导师、统计学教授、研究生同学和当地顾问可能在你选择最适合的统计工具时提供额外的帮助。

补充资源

综合统计教科书

Grimm, L. G., & Yarnold, P. R. (Eds.). (1995). *Reading and understanding multivariate statistics.* Washington, DC: American Psychological Association.

Grimm, L. G., & Yarnold, P. R. (Eds.). (2000). *Reading and understanding MORE multivariate statistics.* Washington, DC: American Psychological Association.

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, NJ: Pearson Prentice-Hall.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole Wadsworth.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison approach* (2nd ed.). Mahwah, NJ: Erlbaum.

Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods* (3rd ed.). Glendale, CA: Pyrczak.

Pedazhur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Orlando, FL: Harcourt Brace.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

Tabachnick, B. F., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.

其他参考

Aiken, L. S., & West, S. G. (1996). *Multiple regression: Testing and interpreting interactions.* Thousnad Oaks,

CA:Sage.

Andrews, F. M., Klem, L., Davidson, T. N., O'Malley, P. M., & Rodgers, W. L. (1981). *A guide for selecting statistical techniques for analyzing social science data* (2nd ed.). Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Bruning, J. L., & Kintz, B. L. (1987). *Computational handbook of statistics* (3rd ed.). Glenview, IL: Scott, Foresman.

Cohen, P., Cohen, J., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (*1999*). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Floyd, F., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 42*, 286-299.

Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, NJ: Erlbaum.

Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology, 65*, 599-610.

Huitema, B. E. (1980). *The analysis of covariance and alternatives*. New York: Wiley.

Jaccard, J., Becker, M. A., & Wood, G. (1984). Pairwise multiple comparison procedures: A review. *Psychological Bulletin, 96*, 589-596.

MacCallum, R. C., Widamon, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-99.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83-104.

Newton, R. R., & Rudestam, K. E. (1999). *Your statistical consultant: Answers to your data analysis questions*. Thousand Oaks, CA: Sage.

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics, 2*, 13-43.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Raykov, T., & Marcoulides, G. M. (2000). *A first course in structural equation modeling*. Mahwah, NJ: Erlbaum.

Singer, J. D., & Willett, J. S. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Thompson, B. (2000). Ten commandments of structural equation modeling. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding MORE multivariate statistics* (pp. 261-283). Washington, DC: American Psychological Association.

Wilkinson, L., & The Task Force on Statistical Inference. (1999). Statistical methods in psychology journals. *American Psychologist, 54*, 594-604.

思考与练习

选择合适的统计方法

早点加强统计知识

—寻找全面的统计书籍、网站和文章

—获得统计分析的经验

分析研究

—识别出需要用统计回答的问题

—列出想要分析的变量

—识别出每个变量的性质(分类、序列、间隔、比率)

—识别出每个问题的自变量和因变量

看看因变量是否适合参数或非参数分析

生成分析计划

—使统计方法与研究问题、假设和变量匹配

—从组比较统计方法中选择

—从相关统计方法中选择

设定 α 水平

假设因果关系时要谨慎

谨慎使用咨询

对自己的统计分析负责

若不确定要寻找帮助

—决策树

—阅读统计书并浏览网站

—寻找与你的设计类似的文章

—请教导师、统计学教授或博学的学生

10 数据的选择、管理与分析

确定研究方案并通过开题委员会的审议之后，就到了收集数据的阶段——数据收集阶段的主要任务就是完成研究方案中计划要做的事情。这一阶段看似只需要按计划行事，但实际操作中却常常会遇到很多问题。虽然详细缜密的设计可以使方案的执行更加顺利，但还是难免遇到问题。本书第3章已经介绍了早期研究计划阶段的相关内容，这一章将向读者介绍如何避免和应对研究实践中可能遇到的问题。

比如：招募被试之前要进行程序检测，以调整好实验程序；训练研究助手（RAs）；做好数据存储和管理的计划；准备好实验仪器和设备等。本章将仔细探讨这些内容。另外需要注意的是，如果你的研究以人类为对象，那么在招募被试之前，就要先取得评估委员会（institutional review board，IRB）关于保护人类（或动物）被试的书面许可。

检测程序

如果到目前为止还没有检测过实验程序,那么,现在就该完成这项工作。正式实验前的程序检测工作非常重要,因为即使准备好的实验程序看似没有问题,也还是可能在被试真正进行实验时出现运行故障。比如:某些被试可能需要用1个小时的时间来完成预计只需要15分钟的程序;仪器和设备的实际运行可能和预期不一致;有些被试可能无法理解指导语等。即使实验程序能够正常运行,也可能会突然发现研究助手人数不够,而临时训练那些不熟悉实验程序的助手是很困难的。

那么,应该选择什么样的被试进行程序检测呢?如果在征集研究有效样本方面没有困难,则可以将前几个被试看作检测被试。此时不要将这些被试的数据纳入到统计中,因为研究者对待前几个被试的方式和对待后来被试的方式通常不同。但如果符合研究标准的有效样本比较少,则应该尽量用和目标人群有最大匹配特征的个体来检测程序,此时建议不用实际研究中真正要用到的被试。例如,在有关儿童的研究中,通常可以请朋友的孩子来做检测被试。有时候还可以请同学、朋友或者研究助手来充当检测程序的被试。

检测程序的目的在于:①确保被试能够作出与指导语相一致的反应;②发现并决定如何处理没有预期到的问题;③测量被试完成任务所需的时间;④学习如何应用和检查设备。然后根据检测情况对实验程序进行相应的调整,解决检测过程中发现的问题。在检测被试完成实验程序之后,可以要求其给出反馈。比如询问被试指导语是否清楚,按指导语要求做出反应时有没有什么困难,实验环境或主试的某些行为是否对其有影响。另外,在检测实验过程中,还要对被试进行观察,看其是否按要求做出反应。然后基于被试的反馈和主试的观察,进行恰当的调整并再次进行检测,再次收集反馈……不断调整直到满意为止。写下注意事项,并提醒研究助手注意这些问题。

如果研究需要通过人为的判断或观察对被试反应进行记录或等级评定,那么在训练研究助手时就要用到检测测试的数据。例如,摄录研究助手将要在操场上记录的活动,或者收集主试将要评估的开放式反应的样本。这样研究助手就可以用真正的数据来进行练习。用真实数据进行练习是很重要的,因为这些真实的数据中常常包含虚拟数据中预期不到的情况。有时候,经过训练的观察者在开始观察真正被试时,观察者内部一致性会有所降低,这可能就是因为训练材料是模拟数据导致的(如,Taplin & Reid,1973)。在收集数据前要先充分训练研究助手,使之能够处理好正式研究时可能遇到的问题。

有时候,还可能通过检测程序发现实验设计或程序存在的严重问题,这时就需要对实验做较大的改动。比如,被试可能会识别精心安插的实验助手,以致对动机研究的实验性控制不能如预期那样操纵被试的动机等。如果变更了实验设计和程序,要得到导师的许可才能继续进行研究。如果所做的改变至关重要,通常需要再次提交到开题委员会进行探讨,至少要提交研究变化的概要,以征得委员会成员的同意。除此之外,也许还需要重新向评估委员会提交申请,获得新的书面许可。

招募并训练助手

进行实验研究通常需要研究助手的协助。例如,可能需要研究助手来做实验主试,或者辅助实验的进行,或者观察被试反应及评定等级等。研究助手可能是本科生、研究生、雇佣者,或者对研究或报酬有兴趣的朋友及亲属。

首先要向研究助手说明工作要求和工作回报,最好是双方签订书面协议。协议内容应该包括研究助手的全部任务以及任务性质,工作的时间和时程及约定的报酬(如果有的话);如果以学分为报酬,就要说明评分规则。同时还要告知实验助手其可能被“炒鱿鱼”的全部原因。

对待潜在的研究助手要坦率、真诚,坦言研究过程中可能出现的非预期情况。比如收集数据所需的时间也许要比预期的时间更长,因为被试可能会迟到或爽约,而且练习测试也需要一些时间。要使潜在的研究助手认识到其工作责任重大,如果不能尽职尽责可能会导致严重的问题。如果潜在研究助手的时间安排与实验要求的时间相冲突,则一开始就要排除掉,这要好于花大量时间培训之后再流失研究助手。因为实际工作量常常要比预期大,这种时间冲突的情况常常会出现。

时间的安排是很重要的。过早招募研究助手是比较常见的失误,这样可能会导致研究助手在随后的几周甚至几个月中无事可做,而当可以真正开始执行实验方案的时候,有些人却不再能参与研究。招募研究助手之前最好确认下列事情:①开题委员会通过了研究方案;②安排好了被试来源;③准备好了培训研究助手所需的材料;④完成了初期的检测程序。开题委员会对研究方案的支持可以确保研究程序不需要做太大的修改,安排好被试来源则可以在培训好研究助手之后马上开始收集数据。

一个研究也许需要几种不同类型的研究助手。比如,一位女性助理、一位治疗师或者一位儿童榜样等。选择研究助手时除了要考虑与实验设计相关的特质外,还要考虑其个人信用情况。他/她是否能够准时并遵守规则?如果潜在的研究助手在职,则可以通过询问其工作情况来了解研究助手的一贯表现。另外还需要考虑研究助手是否具有良好的判断力,最理想的研究助手在遇到没有预期到的问题时会用和研究者一样的方式来应对。例如,要在学校情境中进行观察研究,则需要研究助手着装得体,能够与学校工作人员及老师恰当地进行交流,在教师因事先不知道要进行实验而要求其离开教室时能克制抱怨并保持礼貌等。虽然我们希望研究助手具有较高的智力水平,但很多时候研究助手的信用、社交技能、判断力更为重要。

招募好研究助手后要对其进行培训。培训时可以采用一些有效的训练程序,如“指导语—示范—排练—反馈”的程序:先向研究助手讲解实验任务,再示范如何去做(可以采用角色扮演、录像播放等方式),然后安排其进行练习,最后给出反馈。如果研究助手要完成的任务比较复杂,则可以将任务分解成多个小的部分,一步一步地进行学习和练习,然后汇集起来,这样研究助手就掌握了整项任务。

比如,沙伦·福斯特和他的学生要培训本科生对儿童进行同伴激惹事件的访谈(Dumas et al.,1992)。访谈时研究助手需要掌握一套复杂的编码系统,以此来对儿童的反应打分,同时还要掌握如何控制半结构式的访谈。在培训中,第一步是阅读

编码手册,然后练习对检测实验中的反应进行编码。练习时观察者先从编码简单反应开始,再逐步编码较难反应。研究者事先应该对每种类型的反应进行标准化的编码,培训时将观察者的编码情况和事先确定的标准编码进行比较,并且给予反馈。当观察者编码和标准化编码之间的一致性水平总能够达到 85% 以上时,则可以认为其已经掌握了如何对儿童反应进行编码。研究者还要向观察者讲解并示范各种访谈技巧(如细致地询问和解释儿童语言的意义),然后和观察者一起进行访谈过程的角色扮演,使观察者试着应用所学的技能(在这一过程中要给予反馈)。在掌握了全部的访谈技能之后,可以由研究助手对检测被试进行完整的访谈练习。当研究助手分别对两位不同的检测被试进行了模拟访谈,并且两次访谈都达到所要求的标准之后,就可以开始访谈真正的被试了。

研究助手对实验假设的了解会导致偏见效应(如,Rosenthal,1969),所以要尽量避免使其意识到不同研究条件之间的差别。如果不能保证研究助手对实验条件不知情,至少也要保证实验假设的隐秘性。在避免偏见效应的基础上,还要尽量详细地向研究助手讲解实验,这样才能让他们更好地完成任务。那么如何确定一件事情是否应该告知研究助手呢?可以这样问自己:“知道这件事情是否会对实验助手与被试之间的互动或实验助手对被试的评分产生负性影响呢?”如果答案是肯定的,则不要向研究助手提供这一信息。另外,还要注意避免观察者与实验条件之间的混淆。当在不同的实验条件下安排不同的研究助手时,可能会出现这种混淆情况。这种情况一旦出现,就无法确定研究发现是自变量变化的结果,还是不同研究助手之间存在差异的结果。

当研究助手是为了获得学分而参加实验研究的学生时,要尽量令研究过程对其有教育性。可以通过讲解关于实验设计和实验程序的基本知识来达到这一目的(切记不能透露实验假设!)。通过研究实践中的学习,学生可以了解到标准化程序的重要性、实验控制的概念、观察者内部一致性重要的原因等。研究者还可能成为较低年级学生的良师益友,研究中可以尽可能让研究助手尝试各种不同的实验任务,以此观察和体会研究过程中的各个方面。

遵循伦理规范

假如在研究的准备阶段和申请研究许可的时候曾涉及伦理方面的问题,那么此时则应该向研究助手介绍研究的伦理规范,以使其清楚在研究中应该怎样去做。结合第 7 章研究伦理方面的内容,要记得提醒研究助手保密的重要性。比如不应该同与实验无关的人谈论人类被试的实验数据或行为;不可以在公共场合同其他相关人员讨论被试的情况,因为潜在的被试有可能听到这些评论。

遵循伦理规范不仅要求研究者和研究助手谨慎言行,还要妥善地保存研究数据,这对保障研究的伦理性也十分重要。比如要考虑到如何安全地存储数据,不要在记录表格中体现被试的名字等。又如在学校的情境中对儿童进行研究时,要如何争取教师的同意?如何保证得到每个被试父母的知情同意?如何获取儿童本人的同意?如何保证其在研究过程中有机会提出问题?如何保证被试在不愿意继续参与实验时可以随时退出?

除了完善保密程序和争取知情同意,研究者还要对自己和其他研究助手实施实验的方式进行监控。可以和被试一起听对他们进行访谈的录音,以核实真正获得了被试的知情同意,还要用被试能够理解的语言向其解释他们在研究中享有的权利。另外,实验之后向被试进行解释说明的时候,可以询问被试研究助手是否告知其与伦理原则相关的内容。

另外,如果是对临床样本进行纵向研究,则要特别注意被试临床表现恶化的可能性。如果真的出现了恶化的情况,提交记录被试临床表现的材料会十分有用(Yates,1982)。

设置环境和安排材料

当最终敲定研究程序并安排好研究时间表之后,还要安排其他相关事宜。本章结尾部分的核对表列出了准备开始正式实验前的一系列步骤。这有助于思考如何保障实验程序的运行顺利,并做出相应的计划。包括安排好合适的实验场所,安排在恰当的时间进行实验等。比如要在某一特定的时间段对动物进行系列研究,就要考虑到保证周末和假日里有人负责照看这些动物。

如果研究需要购买特定的量表或者其他受版权保护的材料,还要考虑提前预订(问清折扣和赠品)。同样,如果研究需要实验设备,要确保这些设备在需要时可以使用。另外,还要设计一个备用方案,以防设备突然出现问题。如果有其他研究者要在不同的地方使用这些研究设备,则要互相协商好,保证使用后及时归还等。还要对设备的维护有所规划,比如清洁磁头、更换电池等。

实验前要准备好所有要用到的材料和表格。理想状态是:实验者在被试到达时交给被试一个档案袋,里面按顺序装好所有要用到的表格和准备提供给被试的材料。档案袋里一般包括知情同意书、对实验者的特别提示(如被试的实验分组)以及要被试完成的问卷等。每一页上面都要标明被试的编号,以防这些文件散落。如果实验程序比较复杂,则最好在档案袋里放一张核对表,列出被试需要完成的所有任务。被试完成每项任务以后,由实验者在核对表上做上记号。在被试或者研究助手上交表格等材料时要核对缺失值。例如,在问卷研究中,回收研究材料时要仔细浏览每一项。如果发现有空白项,要询问被试漏下了这一项的原因(如果是有意遗漏,要告诉被试"没关系,不需要对不想回答的问题做出任何答复")。

要妥善保管辛苦收集来的数据。为了保证安全,最起码要将数据存放在有锁的房间里,比较理想的则是存放在有锁的文件柜里。要确保研究助手知道数据存档的规则和程序(如:在一位被试完成实验之后立即将其全部数据存档,不要拿来拿去,原始数据不能随便放置)。存放被试签署的知情同意表时也要谨慎,切不可掉以轻心。每隔一段时间就要对研究数据及资料进行核对,以确保所有被试的全部材料都在。如果被试的真实姓名与其在实验中的编码有联系,则要及时更新并分别锁起来保存。另外,还要考虑火灾或其他可能的突发情况。单独存放数据有重要意义,这值得我们认真思考。

研究者还要考虑如何简便地记录已经参加实验的被试数,随着研究的进行,这一数字会不断增加(张贴显示研究进程的图表可以激励研究者和研究助手),另外

每周还要告知研究助手下一周的研究计划。例如,一位研究者在研究学校情境中儿童的个体间对话时,经常要随着教师的要求、研究助手的时间安排、学生的出席情况及学校教学计划的改变等来调整研究时间计划表。如果不能确定长期稳定不变的时间表,可以考虑在每星期四或者星期五安排好下一周的研究计划。安排研究计划之前要先与学校的教务秘书核对研究场所、教师计划及儿童的时间安排,然后在星期五晚上或者星期一早上将时间表拿到学校,与教师或者相关工作人员协商,确定最终的研究时间表,最后将复印件分别留给教师、学校负责人以及行政管理人员。还要用电子邮箱将确定好的时间表发给研究助手,并在实验室里张贴。在张贴的时间表旁边还要张贴尚未参加过实验且家长同意参加实验的儿童名单。研究助手每天将收集到的数据带回实验室以后,再划掉相对应儿童的名字。

另外一个需要考虑的细节是,如何将数据转化为计算机可以识别的格式。除了不需要进行任何统计的个案研究,其他任何研究都需要这一数据转化过程。有些研究者比较习惯于在收集到全部数据以后再进行数据转换,但最好的方法是采集到数据后立即将其准备好并输入电脑。这样,每个研究助手可以负责输入自己采集到的数据,也要比收齐全部数据后一次性输入电脑容易得多。后面将讨论记录数据的方法。

如果研究者或研究助手需要对录音录像或者记录的反应进行编码,则不适用前面提到过的某些建议。如果对编码方案的执行不够认真,观察者或评价者的个人倾向就很可能引发数据的歪曲。当两个或者更多的观察者使用同一编码系统时,随着时间的推移可能会逐渐形成比较一致的倾向,这时获得的数据就会出现歪曲(Romanczyk et al. ,1973)。这是因为,如果观察者在讨论问题时涉及编码方案,或者形成了编码手册中没有的潜在决策原则,就可能会歪曲数据。如果各观察者的编码倾向性一致,则观察者内部一致性将会很高,并且一般不容易察觉到这种倾向。更重要的是,由于研究初期和研究后期对数据编码的标准常常存在差异,可能会导致出现时程变量、时间系列、前—后设计等的潜在混淆。因此,在这种情况下要先把收集到的全部数据汇集到一起,再对录音录像或者记录的反应进行随机编码,这样可以确保不会出现观察者倾向及数据收集时间的混淆。消除观察者个人倾向的另一个方法是:在研究过程中不断地要求研究助手对标准化刺激进行评分,并根据标准给出反馈(DeMaster, Reid & Twentyman, 1977),这一过程被称为**重新校准**(recalibration)。在需要观察者做出判断的研究中,要格外重视重新校准。如果研究要求观察者在收集数据的同时对之进行编码(相对于后期随机编码),重新校准就更为重要。

对非预期事件做出计划

无论研究方案多么周详,到真正付诸执行时总还会遇到这样那样的问题。要为面对这些非预期事件做好充分的准备,其中一个重要步骤就是确定问题最可能发生的领域。一项研究主要涉及 3 个领域:被试、主试和设备。在这 3 个领域中可能会发生什么样的非预期事件呢?关于被试,有两个因素需要考虑:被试招募和被试守约。招募被试时要记住"三倍原则":想要招募到足够的被试,大概需要预计数量的 3 倍。

有些研究者可能会认为3倍原则并不适用于自己的研究，因为已经有学校同意其在3个班级中进行研究，这3个班级的人数对研究来说绰绰有余。情况可能如预期一样好，但也可能出现意外：如果某班因为教师生病而取消了课程，而另一班级计划进行考试而与研究时间相冲突，应该怎么办？如果班级里愿意参加研究同学的比例比预期小，应该怎么办？如果准备通过打电话来招募被试，那么大概需要打多少通电话才能约到一位被试呢？事先是否能够预期这一比例？

假设现在已经招募了足够的被试，那么是否需要采用一些方法来确保其如期出现呢？这类研究中被试的缺席率一般是多少？如果某被试没有如约出现该怎么办？是重新安排一次实验还是不再邀请这位被试以免再出现类似状况？关于被试守约的问题，可以从牙医的处理方式得到启发，牙医将约定的时间和地点写在一张小卡片上交给病人，甚至可能在约定日期前的一两天打电话提醒病人如期赴约。借鉴到研究中来，我们也可以尝试对被试采用这种办法。

即使我们做出了最大的努力，被试还是有可能会爽约。那么如何利用这段意料之外的时间呢？这是休息、娱乐或者温习功课的好机会么？这种想法是不对的。这段时间本来是计划进行研究工作的，即使被试没有出现，也并不是真正的空闲时间。如果打算按照原计划进行研究，就要在以后找机会来补上这段研究，这将会占用计划进行其他相关研究的时间。我们建议，要事先对被试不能守约的情况做出安排，在此期间做一些与研究相关的工作。比如可以修改论文的引言部分，为最终的毕业论文或者发表研究结果做好准备；一般在这个时候已经进行了一部分实验，因此可以进一步完善研究方法部分；也可以将已经获得的数据输入电脑；或者查找一些不易获得的文献资料等。这样可以充分利用被试爽约所空闲出来的时间，而且不会偏离完成研究目标的路线。

出现非预期事件的另一个可能领域在于研究助手或者其他研究相关人员。人人都可能遇到生病等意料之外的情况，以致不能如期赴约，还有可能忘记收集或录入一些重要的数据等。那么应该如何处理这些意外情况呢？最好是采用标准化的人员管理方案。首先要选择最合适的人当研究助手。当然这也许还和报酬等因素有关，如果能够提供适量的金钱或者学分作为报酬，则可能请到更合适的研究助手，并从中有更多的收获。如果不能付钱，还可以考虑其他的回报方式。比如可以采用“互相帮助”的方式，帮助研究助手编写实验程序或分析数据等，或者可以每采集到几个被试的数据就请研究助手吃顿饭，也是不错的主意。

为了避免因为研究助手请病假或者休假带来问题，可以对其进行交叉训练①，即保证每个研究助手至少了解另外一个人的研究工作，这样就可以在必要时找到替补。

另外还要计划好对研究的监控。除了要监控研究的伦理性程序，核对收集到的数据，还要确保研究助手严格遵守实验程序。再认真的研究助手也有可能出现失误，定时核查其表现有助于在问题失控前发现它。

---

① 交叉训练是指训练一名研究助手掌握另一名研究助手的工作。比如，一项研究中可能需要调试设备的研究助手和观察被试的研究助手，那么训练这两位助手互相了解对方的工作，这样当调试设备的助手临时有事不能如约参加研究工作时，就可以由观察被试的研究助手来调试设备。——译者注

最后,设备方面也可能会出现意外,需要对此做出周全的准备。某些问题(比如,投影仪的灯泡烧坏了,投影仪和计算机连接不上等)会比另一些问题(如,测谎仪被盗)更为常见,但是最好是事先想好各种有可能出现的意外情况,并做出周密的准备。如果研究者认为自己进行的是技术性较低或无需技术的研究,而且不需要任何设备,则需要重新仔细思考是否真的是这样。比如,想一想去研究地点所乘坐的交通工具,又如用来进行词语加工或数据分析的计算机。曾有一位研究者一年中在不同的电脑上烧坏3个存储数据资料的硬盘,这警示我们应考虑同时在几台不同电脑上备份文件和数据。

总的来说,可以采用一些比较好的办法来处理这3类潜在的意外。其中对研究助手进行交叉训练是非常重要的一个办法,另外还要准备充足的实验被试、研究助手和研究设备,事先计划好研究设备出现问题时要去哪里租借设备应急,预先安排好替补研究助手等。美国航空航天局(NASA)在其外太空载人航天研究中就特别强调备用方案,如果一套方案失败了,马上就能够启动替补方案。以这种方式进行实验研究,以这种态度对待生活,将不会受到意外事件的影响和困扰!

收集数据

现在,做好了一切准备工作,我们终于可接待第一位被试了,或者终于可以开始对第一个动物进行正式实验了。假设程序运行很顺利,被试完成了整个实验系列,每个人都很准时,并且各尽其职。那么,是否就可以在研究助手采集数据时放松休息或者娱乐一下了?当然不能!

保证整个实验过程顺利进行的一个关键因素是**严密的监督**。虽然这不意味着事必躬亲地关心研究助手的每一个细节,但是一定要确保助手们及时并正确地完成其职责。为了做到这一点,通常要设立目标并为自己和研究助手做好计划,然后按照计划一步步地执行,直到完成研究任务。当研究助手表现好时要给予积极的评价以及其他正性反馈。同样,如果研究助手存在不足也要及时并礼貌地与之交流,以防在没有监督的情况下出现问题。同时要鼓励研究助手遇到问题时马上拿出来探讨,无论是程序、数据编码还是其他方面的问题,都要及时发现、及时解决。还有,要定时召开研究小组成员会议,讨论研究进展和其他相关问题。

另外,在拿到数据时还要进行核查,看收集到的数据是否存在问题。如果是在研究助手在场时进行这项工作,注意不要对较好的数据表现出过分的赞扬,这样也许会给没有做到这点的助手带来消极影响。这种行为可能会促使研究助手为了取悦研究者而使数据出现偏向,不能真实反映被试的实际表现(O'Leary, Kent & Kanowitz, 1975)。

如果需要由不同的人来处理和加工数据,就要考虑追踪每一部分工作由谁完成。假设要收集、记录一组基于访谈的磁带录音数据,并对其评分。那么一收集到这些数据就要立即对其进行标记,然后做访谈记录。保存好这些记录,之后依据其进行评分时还要核对。那么应该如何设计和组织追踪程序呢?一个办法是使用如Excel或者Calc之类的电子数据表,或者还可以采用技术含量比较低的手写核对法。

如果是在应用情境中收集数据，还要定期与情境中的相关人员进行核查，以确保一切运行顺利。这样做可以及早发现程序中潜在的困难和问题，这是至关重要的：如果有一名研究助手没有遵守研究规范，也许就会影响到这一情境中的其他数据，以至全部数据都不能使用。研究计划发生变更时要告知情境中的关键人物，完成数据收集以后还要正式对其表示感谢（可以私下表示感谢，也可以通过张贴感谢信等方式来表示）。如果之前承诺过向其反馈研究结果，就要在数据统计分析之后反馈研究结果。一名研究者的表现常常会影响到后续研究人员是否能够获许在此应用情境中进行研究。①

最后，要在研究日志中记录数据收集阶段发生的比较特别的事，比如研究者的决策、非正式的观察等。简要记下改进研究的初步想法，这可能在以后写作论文的讨论部分时凸显其作用，尤其是在研究结果不同寻常时会更为有用。在统计分析数据时考虑数据采集时的情况，也许有助于对研究发现做出更合理的解释。

存储、核对和分析数据

当准备开始收集数据时，就要考虑好如何存储数据，并将之转化为计算机能够分析的格式。一条重要的原则是：从原始数据到计算机识别数据之间的步骤越少越好。一个不错的方法是由被试直接使用电脑做出反应，同时记录其答案。这需要准备好计算机设备和测量程序，且研究中采用的测量要适用于这种方法，被试要有效地掌握计算机相关技能（或受过计算机使用训练），这样才能保证程序可以进行。另外，如果测量需要以纸笔的形式进行，那么在将数据录入计算机时要确保准确无误，同时保证数据可信并且有效地反应测量的情况。如不方便直接用计算机采集数据，可以考虑以计算机能够读取的形式来采集数据。这样就可以直接进行统计分析，而无需再手动录入数据。如果不能直接用计算机采集数据，也不能以计算机能读取的形式采集数据，那么研究者和研究助手就要人工将问卷或成绩单上的数据录入到电脑中，注意要尽量减少中间步骤（如将数据抄到一张表单上再进行录入）。

在开始录入数据前，要先制定数据录入指导手册。包括计算机能够识别的变量名（如，在统计分析时用到的变量名称），对变量名所对应数据的描述，对变量水平的划分。例如，编码手册也许会要求在计算机上对一个种族问题输入反应，对应在变量名为“种族”的统计表格下，是非裔美国人则输入“1”，是亚裔美国人则输入“2”，等等。

在将数据转换成计算机可读的形式时，要考虑到所用的计算机程序。每个程序对数据的格式都有不同的要求。如果是对单个被试跨时间的重复测量，就需要考虑将要使用的程序中对重复测量数据的格式有怎样的要求。有些程序要求将数据分行，有些程序则要求同一被试的所有数据都在同一行上。

此外，还要考虑将要使用的程序如何处理缺失值。处理缺失值要求删除含有缺失值的一列或者一对数据，具体来说，常用方法有两种：一个是删除有缺失值被试的

① 如果一名研究者在某学校进行研究时给该学校留下了良好的印象，那么该学校则更可能同意其他研究者来进行研究；而不良的印象可能导致学校不再接受其他研究者的研究申请。——译者注

全部数据;另一个是,如果成对数据中有一个是缺失的,则删除这一对数据。例如:如果80个被试中缺失了两个被试的放松训练时间,则对放松训练时间与其他变量的相关统计就要基于78个而非80个被试进行的。在对放松训练时间以外的其他变量进行统计分析时,则还是会包括这两个被试的数据。本章的后面将会有一部分专门讨论与缺失值有关的问题。

如果对计算机或将用到的程序不够熟悉,可以和一些比较有经验的人交流,学习将会用到的数据分析方法以及所需要的实验数据格式等。对一些比较常用的程序,常常可以找到其指导手册(如,Pallant,2001,for Spss;Delwiche & Slaughter,2003,for SAS)。大部分程序都有其基础的使用说明,包括如何建立数据文件、数据管理(如,变量转换)、解释输出结果以及一些统计相关的基础内容。

如果想自己购买程序软件,则要小心选购。应结合将进行的数据分析来考虑,在购买之前问清相关问题。一些软件公司,比如SPSS,可以在线免费试用,看其功能是否与所需的分析相匹配。如果要进行的分析十分特殊或者非常重要,那么事先的考察就尤为重要,有些大的软件包可能也不能满足这些数据的统计需求。要考虑研究生的经济能力,以及需要买程序的完全版还是部分相关功能(如,只处理有限数量的样本和变量)。

有时候可以在线使用一些免费软件。如果可以下载到符合统计要求的相关软件,就不需要再买标准软件包了。这些免费软件是由一些有能力、有理想的专家开发的,他们致力于建立免费的、公用的统计软件。例如:Mx(http://www.vcu.edu/mx),是建立结构方程模型的一种程序语言。从经验的角度看,这种可以免费获取的程序大多是适合非常熟悉计算机编程的人使用的,可以用于解决统计分析中相关的程式问题。另外,因为免费程序的开发者需要为之付出时间和精力,所以这些程序通常不能进行太复杂的统计分析。对于统计专家和计算机专家来说,这不成为问题;但对新手来说,最好还是选用较为简便通用的软件包。

在计算机文件中应该记录哪些数据呢?最好是记录和分析有可能相关的全部数据。例如:尽管在数据分析时很少用到主试这一变量,但是研究中还是经常要对此进行记录(这些资料还是有用到的可能性,比如当数据比较奇怪时,也许需要找出是哪位主试记录的数据)。如果想要计算机给出次级量表,或者想计算问卷的内部一致性,或者想以不同方式重新分组,则要对问卷项目单独编码。如果没有这些要求,则可以考虑只记录总分。记录数据的一个原则是要尽量详尽地记录数据。后期如果想要对数据求和很容易,但是如果想分解一个总分却很难实现。①

真实地记录数据,需要在人工记录的每一步中都加倍小心。在数据输入阶段,一般可以采取他人核对的方式保证其准确性。一个简单有效的方法是成对地工作:一个人读数据,一个人输入数据,然后两个人互换来核对。一些统计分析程序可以对同一组数据进行双重输入,当发现两次输入不符时会做出提示。也可以对每一变量进行频次统计,通过频次统计来找到特异值(比如1—5的五级量表中的42)。

核对不仅限于数据输入部分。如果是通过手动求和来给问卷计分,就要请其他

---

① 比如在问卷研究中,如果记录了每个项目的得分,很容易就能得到总分;但如果只记录了总分,就无法了解被试对每个项目的反应,不能进行进一步的分析。因此,要尽可能详尽地记录数据。——译者注

人进行检查。如果通过变量转换程序来计算分量表得分,最好手动计算一部分得分来和计算机的结果相比较,以确保转换程序的正确性。核对的过程十分重要,一定要认真完成。研究者和其他人要分别检查和记录计算过程中的主要错误。这些错误会影响甚至改变研究结果和对结果的解释。切记,计算机无法分辨出数据的记录是否正确,给它什么数据,它就分析什么数据。如果输入的数据是无效的,那么输出的结果也就是无效的,所以,要确保数据输入的准确性。

在记录和核对数据之后,就可以开始进行分析工作了。首先要列出研究计划中计划进行的分析,第 9 章已经就这部分内容进行了讨论。在检验假设之前要先进行初步的分析。如第 9 章中提到的那样,初步分析一般包括以下内容:①核对检验假设需要进行的分析计划;②检验测量的真实性和有效性;③检查非故意混淆;④检查方法学上的细节,比如顺序问题等,是否需要解决;⑤检验变量的分布;⑥检查数据缺失情况,以及缺失值是否与研究中的其他变量相关。初步分析的结果可以帮助我们了解是否适合按计划来检验假设,是否需要重新思考假设检验的统计程序。

先花一些时间观察数据。如果研究考察了多组被试,要画出变量间相关关系的散点图。看每组被试在主要变量上的分布情况,不同组之间的分布是否有重叠;如果有,重叠部分有多大。还要看变量的频次分布。通过对数据的前期观察可以得到对数据的初步印象,这对进行后续的复杂统计分析可能有所帮助。

在进行组间设计的主要分析之前,要确定各组在统计特征以及其他潜在混淆变量上是否一致。比如在一项研究中,要计算人口学变量与独立或者预测性变量之间的相关。如果人口学变量(或其他潜在混淆变量)与自变量存在关联,则可以计算潜在混淆因素和因变量之间的相关。如果混淆变量和因变量相关不显著,则潜在混淆因素与因变量不是线性关系,不会影响分析。而显著的相关则表明,这一混淆变量是需要控制的。或者通过设计变更(例如,使用混淆变量作为额外的自变量),或者通过统计学方法(例如,利用协方差分析,见第 9 章的讨论)来进行控制。

此外,还要检查数据是否符合统计分析的要求。如果不符合要求,则需考虑统计计划是否符合研究假设。如果不符合假设,就要采取适当的措施来调整(如,进行数据转换,或选择其他分析方案)。可以参考第 9 章中对各种数据对应统计方法的讨论。

最后,要寻找并处理缺失值。当将数据输入到计算机时,总是会发现人工的操作存在问题,甚至可能发现缺少某些被试的全部数据。那么应该如何处理这些缺失值呢?

这个问题看似简单,但实际回答起来相当复杂。对这一问题的处理取决于数据最初缺失的原因。数据的缺失可能有 3 种原因:①与自变量和因变量都有关("非忽略性缺失" MNAR);②与自变量有关,但与因变量无关("随机缺失",MAR);③与自变量和因变量都没有关系("完全随机缺失",MCAR)。谢弗和格雷厄姆(Schafer & Graham,2002)提出了在这些不同情况下处理缺失值的常见方案,这里主要参考了他们的分析,并强烈建议研究者通过自己的理解和讨论来深入探究这个问题。

通过预分析,可以知道缺失数据是 MCAR、MAR 还是 MNAR,接下来有几种处理缺失值的不同方法可以选择。其中最常见的是整行删除,如前文提到过的。这时任何数据的缺失都会致使这一被试的全部数据不可用。但在实际应用中这种方法

的缺点很明显,整行删除会导致样本数大量减少。此外,除非该数据的缺失是完全随机的,否则整行删除经常会导致结果出现偏见效应(Schafer & Graham,2002 年)。但如果缺失数据比较少,还是可以考虑选择这种方法的。

处理缺失值的另一个常见方法用于应对项目或多项表反应缺失的情况。这种方法是用其他样本的平均值来填补缺失值,例如当一个或多个项目缺失时,可利用这一项目非缺失项的平均值作为缺失项的值,称为非加权平均估测(Winer, Brown & Michels,1991)。谢弗和格雷厄姆(Schafer & Graham,2002)报告,这种方法也可能引发偏见效应,但当其他项目内部相关性较高(即内部一致性较高)时,偏见效应相对较小。另外还有些方法涉及用回归方程来估算并重置缺失值(Farrell,1999)。

以一个替代值来取代缺失值也是一种估算方法,这要求对缺失值做出较好的猜测。还有一种处理缺失值的方法需要大量的计算,为缺失的部分生成一些可能的值,然后对这些可能值进行平均。虽然通常认为多值平均估算方法优于单一估算方法(Schafer & Graham,2002),但这些新方法还没有得到广泛应用。最后,还可使用最大似然法①,这些统计技术利用现有数据,对被试的缺失值做出最佳评估。但是上述方法只适用于随机缺失(或完全随机缺失)的数据。这些方法都可以通过许多常用的统计软件包来实现。

学习应用统计软件

进行统计分析不仅需要了解数据,还要对所用软件有所把握。如果是第一次使用某种软件或初次进行某种分析,要做好心理准备:结果可能不如预想的好。一些计算机程序附带帮助菜单,但是通常只是说明主要技术问题,缺少对常见数据处理的介绍。每位研究者都有可能会遇到问题,而且计算机并不总是按照使用者的意愿行事。那么在研究初期就应该学习并掌握一些数据分析程序,要正确应用一项新程序进行数据分析大约需要 7 次尝试。

如果对即将进行的统计不熟悉,可以先向比较了解这一软件和统计的人请求帮助,至少请求其给出一个应用该统计程序的例子。比如可以向导师或者是研究小组中的其他人寻求帮助,并选择在方便获得帮助的时候进行分析。分析之前最好做好准备,因为统计分析的顾问可能希望看到已经得到的输出结果和报错信息。

在告诉计算机做什么的同时,还要花时间来告诉它如何为打印输出以及变量注明标签,给出明确的名称。标签可以标识自变量的意义及其水平,变量 A 或 B 这种名称在当时可能意义明确,但一个月以后对其含义的记忆就开始模糊不清了。标签命名也许需要一些时间和想象力,但编写程序可以使每页顶部呈现项目名称及自变量和因变量名称,这将会使以后的操作十分便利。如果你有过盯着电脑努力思考几年前创建的文件名的含义的这种经历,则会非常赞同这一建议。

如果统计程序无需编程,只需通过鼠标点选(例如 SPSS),则最好学习使用和保存语法文件。计算机语法是告诉计算机如何工作的语言。这种语言是隐蔽的,在使用程序时看不到。但是,我们要学会取得所使用程序的语法,将其保存、记忆并重新

① 一种非线性的滤波方法。——译者注

利用。如果能在使用后将之保存在计算机中,在重新分析数据时则只需花费很小的精力。

语法文件为什么重要呢? 想想“通过7次尝试才能做得正确”这一经验。假设你已经为所有的变量计算了总分,然后发现其中一些分数很奇怪。如果保存了语法文件,就可以回头检查是哪里出了差错;如果没有保存,则只能对数据完全重新分析。假设后期核查时发现遗漏了一个被试的数据,就需要重新进行所有的分析;或者想检查过滤掉的一些被试数据是否有可能出现问题,是否与原来的结果一致。应用语法文件,重新分析只需要几分钟,而如果没有保存语法文件,就需要从头开始逐步重新分析,这可能需要几个小时。

此外,要考虑数据分析程序的预设值。如果不主动设置预设值,计算机就会自动设置。注意,数据分析越复杂,需要选择的统计分析算法就越多。例如,因素分析包括有关如何提取因素的选择等。不同的统计分析需要进行不同的选择,所以每一个分析软件都可能有不同的默认值。计算机程序员努力尝试在各种程序中使用广泛接受的或常见的默认值,统计学家也讨论过各种默认值的优点和缺点,然而程序默认值并非永远都是最佳选择。是研究者,而非 SPSS 或 SAS 的程序创造者,要对统计的细节以及默认值是否适用于自己的研究负责。大多数程序允许重新设置默认值,因此最好是根据研究数据选择最合适的统计值。

核查数据和初步分析结果之后,还要考虑是否需要修改检验-假设的计划。如果分析遇到困难或者出现统计方法学上的问题,可以与导师讨论。不要根据自己的主观偏好选择统计分析方法,最好是在做研究计划时就确定采用何种软件进行哪些统计分析。

当进行假设-检验的分析时,要知道计算机统计样本的大小和自由度,并检查输出,看计算机给出的值是否与之相符。如果不相符,就要找出程序或数据输入的错误之处。可以请导师或其他更有经验的人来检查输出的结果,以判断问题之所在。同时还要记录数据的异常之处以及程序设置情况,以便后期重新思考(例如,数据分析时剔除的特异值、变量的转换等)。

最后,要计算效应大小和置信区间(Wilkinson & The Task Force on Statistical Inference,1990)。对于不显著的结果,要检查统计力。如果设置要求报告这方面的信息,则大多数计算机程序会自动进行检查,其效应大小有助于解释研究发现的可推广程度。一些严谨的统计学家认为这远比 P 值是否低于设定的 α 值(如,0.05)更重要。统计检验力可以对非显著性的结果做出解释,计算统计区间则有助于将研究结果与其他研究结果相对比。

补充性的探索性分析

如果在分析数据时发现需要进行一些额外的分析,而这些分析超出了最初的计划,不要惊讶。近期的一项研究结果证明,严格遵照预期进行分析,常常会遇到一些问题。额外的分析经常有助于解决这样的问题,而且有助于更全面地理解数据。这种探索性分析的结果不一定影响最终的结论,是否有影响取决于分析结果的意义以及计划对结果进行的解释。在研究后期撰写结果和讨论部分时,也许还会发现需要

进行一些附加的分析——在后面两章中将会谈到这个问题。

补充资源

统计软件指导

下面提供 SPSS 和 SAS 的新手指导。可以通过相关综述或图书零售网站上的介绍来获得更多资源(如,Amazon. com,Barnesandnoble. com)。

Delwiche,L. D. ,& Slaughter,S. J. (2003). *The little SAS book*(3rd ed.). Cary,NC:SAS.

Green,S. B. ,& Salkind,N. J. (2004). *Using SPSS for windows and Macintosh:Analyzing and understanding data*(4th ed.). Upper Saddle River,NJ:Prentice-Hall.

Pallant,J. (2001). SPSS *survival manual:A step-by-step guide to data analysis using SPSS*. Philadelphia:Open University Press.

UCLA Academic Technology Services. *Statistical computing resources*. Retrieved November 15,2005,from http://www. ats. ucla. edu/stat. [UCLA Statistics Department Web site;has some great information on SAS, SPSS,and STATA basics.]

缺失值

Schafer,J. L. ,& Graham,J. W. (2002). Missing data:Our view of the state of the art. *Psychological Methods*, 7,147-177.

思考与练习

数据的选择、管理与分析

☐在写作方面获得 IRB 的赞同

☐确定实验材料和研究程序

—开发新的刺激材料

—重新思考控制条件

—检测程序

—在导师或研究小组的指导下修改研究程序

☐安排研究条件和设备

—获得研究场所

—安排研究场所

—获得学习设备

—学习设备使用规则

—获得数据存储设备

—为数据分析准备合适的计算机和软件

—安排实验助手、被试、房间、仪器设备

—安排假期和周末参与研究的被试

☐安排被试招募

思考与练习

—确定被试参与研究的时间表
—建立被试记录表
—开始招募研究被试

□准备好表格和跟踪程序
—复印知情同意表
—复印其他表格
—建立对被试进行追踪研究的系统
—确定筛选数据和被试信息的系统

□招募和训练研究助手

□对非预期的事件做准备
—被试
—研究助手
—仪器设备

□收集数据
—安排与助手见面的时间表
—监督助手的表现
—及早发现可能的问题

□记录数据
—核对数据

□输入数据
—建立数据录入说明
—建立数据库
—确定数据核对计划
—数据输入
—查出并核对特异值
—检查数据转换程序的准确性

□分析数据
—如果不熟悉计算机程序,要寻求帮助
—考察默认值,设置适合自己统计分析需求的程序
—回顾数据分析计划
—进行最初的分析
—考察缺失值,并确定如何进行处理
—进行假设-检验分析
—进行探索性分析
—计算效应大小、效应力和置信区间
—为变量和输出结果注明标签

11

报告研究结果

收集到全部研究数据之后要对其进行分析(初步的分析往往并不能尽如人意,需要在反复思考、多次试验之后才能进行正确的分析),然后报告研究结果。一般情况下呈现研究结果时不需要报告采集到的原始数据,研究者更多地倾向于直接报告对数据的分析结果。报告研究结果之后再呈现研究的讨论部分。

如果前面的每一步都按照指导进行,并且一切顺利,那么论文的结果部分就易于撰写,因为前面的研究早已为此做好了充分的准备。很多人对前面章节中关于模拟预期结果的建议不屑一顾,研究进行到这一阶段就可以验证这一建议的重要性了。模拟预期结果有助于确定实验设计和统计分析计划,还提供了在结果部分进行分析说明的框架。事实上,如果遵循研究建议,很好地完成数据模拟这一前期准备工作,那么就已经完成了分析研究数据的大部分工作。在这种情况下,研究者的主要工作就是筛选数据,并用真正的研究数据来取代模拟数据。下面将介绍论文结果部分应该包含哪些信息,以及应该如何组织这些信息。

首先应该将结果部分的内容限制在与研究目的相关的分析上,要避免呈现有趣但是偏离研究目的的内容。许多学生或者研究者不遵循这项建议,以至于最后不得不重新分析处理与主要假设相关的数据。例如,为了确保研究足够充分,一项研究可能会以一种或几种形式来收集各方面的信息(比如第 8 章中讨论研究信度和效度的例子)。假如研究中的主要因变量是由他人评分采集的,则需要在验证主要假设时进行额外的分析,比较相互独立的不同评分者所评出的分数,以证明评分是可信的。虽然不同研究者间相比较的结果与主要研究假设并没有直接的关系,除非研究的目的就在于考察某一测量方法的内部信度,但是呈现这些信息依然是十分重要的。上述数据最好呈现在介绍测量方法的方法部分,也可以把与样本特征相关的数据及各组的人口学信息放入论文的被试部分。但是要注意,也许会有人不同意这一观点,因此要与导师认真讨论。解决了统计分析方法学上的问题之后,就可以将注意力集中到主要数据上了。

对数据进行基本处理

将报告研究结果与请客吃饭相类比也许会更简明易懂。假想那些主要数据就是客人,论文中将呈现这些数字以支持假设,那么就需要做一些家务来准备招待这些客人。首先要确保具备晚餐所需的一切材料,我们一般称之为复杂的、合格的、有限的(CQL)结果部分。这包括对缺失值、被试流失以及许多可能影响样本代表性的问题的处理。在第10章中已经探讨了处理缺失值的方法。结果的CQL部分需要介绍是否有数据缺失,缺失了多少数据,为何出现缺失以及是如何处理的。如果有被试中途流失,要说明被试所在组以及是否会对组间的对比产生影响。当不同组间有不同的被试流失或增加时,会对结果产生影响。这一部分还要说明其他可能会影响结果的偶然因素。

准备好举行一次成功的晚宴所需的全部原料之后,还要花点时间检查每一种材料的质量。是用真正的奶油还是用人造黄油?酒怎么样?花点时间对数据进行批判性的思考。收集到的数据是否存在明显的错误?是否需要在进行统计分析之前处理这些问题?这并非是指修改或删除不支持假设的数据,就像威尔金森统计推理团队(Wilknson & The Task Force on Statistical Inference,1999)所指出的那样,“数据审查并非数据删选”(597页)。例如,观察数据的分布,能够知道是否需要对数据进行转换,使之适合于研究方案中计划进行的数据分析。

这一部分还要报告数据是否符合将要进行的统计分析的要求,介绍采用此种统计方法的原因、步骤和结果,以及根据输出结果对统计分析进行的变更。尽管在较长的实践过程中一直使用统计的方法检验数据是否符合统计假设,但美国心理学会(APA)提出了统计相关的任务需求(Wilkinson et al.,1999),认为图表分析更为合适。这是一项很好的建议。此外,在进行数据分析前总体考察原始数据的质量也很重要。

报告初步分析结果

现在客人到了,已经用过了餐前甜点,每个人都在期待主菜,即:将要呈现的与检验研究假设相关的主要数据,以及证明结果有意义的统计处理方法。同时还要考虑平均数和数据的偏态。

如果在介绍结果时能够有逻辑、有条理,那么论文会显得十分简洁。除非进行个案研究,否则要避免介绍个体分数和原始数据。最好保留原始数据,如果需要的话可以把它放在附录中。正如耶茨(Yates,1982)所说,以简单的表格保留原始数据非常重要,尤其是在投稿后编辑要求进行进一步分析的时候,这些数据更显得尤为重要。结果部分只要说明数据及对数据的处理方式就可以了。至于这样的结果说明了什么,要留到讨论部分去说明。

值得高兴的是关于结果的介绍没什么神秘也没什么难度。一旦决定了这部分要包含什么、避免什么,剩下的就是程式化的内容了。

列出研究结果

要想使研究结果更有条理,需要重新回想到构想假设的阶段,回想与研究假设有关的那一部分内容。如果在着手进行研究之前已经提出了一些假设(如第 4 章介绍的),那么最好根据重要性将之排序。这样同时还可以对论文,尤其是结果部分的书面组织有所帮助。介绍主要数据分析结果时,要按照这些结果对研究假设的重要性来排序,给出检验每一条假设的数据及相关的统计分析。如果在假设中有较多的因变量或较多的分析,可以考虑用小标题将其标出。每个小标题都应简短地描述其所检验的假设。一个全面的策略是先陈述研究结论,然后附上相关数据和统计分析来支持这一结论。例如,可以说"研究发现男生在大运动量活动中花费更多的时间,这从表 11.1 中可以看出,在全部 3 项大运动量的活动中,男生的参与程度都显著高于女生(平均性别效应[1,59] $=4.62, p<0.05$)。"一些权威人士建议根据主要因变量的测量去组织研究结果。这样似乎没有根据研究假设来组织结果合乎逻辑,因为研究目的并不在于测量本身。相反,应该把测量作为检验假设的一项工具。如果存在一些和假设相关的测量,那么可以在假设中用小标题来对其进行组织(如,被试自己陈述的数据,被试父母报告的数据)。即使没有使用小标题,也一定要有序地在每项测量中呈现结果,然后在呈现其他假设的相关结果时也遵循相同的顺序。

表 11.1

暴露和非暴露模式下不同性别的自我暴露水平:平均数和标准差

	女性		男性	
模式	*M*	*SD*	*M*	*SD*
暴露	7.60	2.34	4.85	2.02
非暴露	4.76	1.89	2.67	1.93

注:$n=22$ 人/组

有时根据分析来组织结果也可以十分清晰。比如有这样的 3 个假设,它们是关于多元回归方程式中 3 个变量的不同百分比。在这种情况下,如果想总结所有的假设,那么就可以只呈现多元回归的结果。这种方法避免了冗余,而且能保持结果和假设的一致性。

报告相关信息

对于所报告的每一项分析,都要包含其统计的名称及属性。例如 ANONA 的属性,包括因素、每个因素的水平、是否有些因素存在重复测量,还有所用到的事后检验的名称等。要包含所有显著效应的相关统计值(如,F, t, p)。要报告研究选择的

统计显著区间的 p 值，并说明选择的依据。一些人也许希望研究中提供非显著性的效应，那么可以将之呈现在附录中。还要报告平均数、标准差以及每种条件下的样本量。对相关研究，应该把这些样本作为整体，除非需要对样本进行单独分析（当需要平均数、样本大小和每个样本组的标准差时）。另外，需要强调的是要报告分析的效应力的大小。这是因为 p 值并没有说明相关的强度或者效应大小（见 Wilkinson et al. ,1999；APA,2001）。此外，还要报告测验统计的置信区间（如，95% 的置信区间[CI] = 1.08 ~ 2.77）。

如果进行组间比较，需要报告每项显著效应的平均数、标准差及样本量。例如，如果一个性别与年级的 ANOVA 分析仅仅显示出现了性别主效应，就要报告性别相关的统计量（而非年级相关统计量）。也可以在附录中报告每种处理的相关统计量。文章中的表格或者图表应该包含哪些信息呢？一般应包括平均数、标准差。如果数据很少的话，可以不用表格，直接在行文中呈现。在表格或者文章中都可以报告统计值（如，F_S），作者可以便宜行事。另外，需要说明的是，在文章中需要说明统计检测的名称及相关属性。

清晰地表达研究结果

切记，结果部分需要最清楚的表达。这不是追求创新或者追求语言美的时候。"简单地重复"是一条需要遵循的重要法则。仔细考虑，然后选用表达最清楚的特定句子结构，并且采用这种句子结构陈述所有研究结果。假设一项研究采用的是 2×2 的因素设计，包括 3 种不同因变量，使用 ANOVA 进行分析，那么就应该以相同的方式报告每个因变量的统计结果。如果第一个自变量为性别，就最好在报告每一结果时都首先呈现它，接下来才是第二个变量，变量之间如果存在交互作用则报告交互作用。如果在报告第一个因变量时先呈现是女性的数据，那在报告其他因变量时也先呈现女性的数据。一致性与对称性会使研究结果更易于理解。应尽量减少句型的多样性，简明扼要的说明会使读者更容易理解论文内容。

而且，一致性的报告方式便于读者回溯以前阅读过的内容，以更好地理解或者将前后的内容相比较。除了单一重复，直接用语言表述会比写统计句子更合适。要养成将统计数据放在句尾的习惯。例如，"男性比女性更不容易注意到表情的差异（$M_{males} = 5.2$；$M_{females} = 3.5$），$F(1,29) = 6.49, p < 0.05$"要比"ANOVA 分析显示 $F(1,29) = 6.49, p < 0.05$，男性在试验中注意到表情变化的概率是 5.2，显著高于女性的 3.5"清楚得多。

注意，在第一个例子当中，在比较男性和女性时同样讨论了因变量，却没有涉及因变量的测量。在对被试变量进行单一测量的时候这是可以接受的，但如果进行了不只一项测量，则可以使用以下方式来处理："在面部表情测试中，男性比女性不易注意到面部表情的差异。"然后报告第二项测量的结果，例如："但是，在使用录影带进行的测试中则无显著差异（$M_{males} = 12.4$；$M_{females} = 13.1$），$F(1,29) = 1.56, p = 0.14$。"

统计报告的相关要求

通常，报告数据统计结果的时候需要遵守一些规则，比如坚持"简单地重复"原则有助于读者的理解。遵守另一些呈现数据的要求也有助于理解文章，使读者在弄

清报告数据的方式之后不会在理解研究发现时感到迷惑。《美国心理学会写作手册》(*Publication Manual of the American Psychological Association*)中说明了心理学方面(特别是社会和行为科学)的论文报告研究结果时需要遵循的规则(APA,2001)。前文中已经给出了一些根据APA规则呈现的研究结果。在论文中报告统计信息时,最基本的规则就是先给出抽样的自由度,然后报告统计值,最后是可能性。小数点前面一般不会出现单独的零。当进行卡方检验时,要在括号中呈现自由度和样本大小,例如,$\chi^2(4, N = 86) = 9.67, p < 0.05$。这里不再对统计数据的报告方式进行详尽的介绍,《APA写作手册》对这方面的内容有更详细的解释。

讨论补充分析

涉及主要假设的分析经常会引出进一步的问题,从这里着手可以进行更多的深入分析。如果某一项分析有数据支持并且与论文主旨相关,则建议在论文的结尾部分附上这一补充分析,这将成为论文的重要补充部分。通常情况下,这部分是为进行大胆的思维推理所准备的,还可以引领研究者进一步探索论文研究的问题。然而,不要让这种新异的兴趣点导致研究偏离当前的主要假设和任务。记住,尽管新奇的发现可能会很重要,但是不要让其影响到主要假设,因为这些假设是建立在对该领域研究的综述和一系列认真研究的基础之上的。

编制表格

表格很难排版,因此,除非重要的内容无法通过文字清晰地叙述出来,一般情况下大部分的杂志编者都会抵制表格。而学位论文一般不需要对表格进行复杂的排版,因此具有更多的选择性。

精心设计的表格在帮助读者理解研究成果方面有不可估量的作用,特别是当这些结果很复杂的时候。但是,这并不仅仅需要在计算机软件中剪切表格,再粘贴到论文中去。这种自动输出的表格一般都不是标准的格式或尺寸,因此也不符合论文要求。熟练地建立表格来呈现数据需要大量的练习,所以采用这种简单的呈现方式时却往往很麻烦。本书能给出的最好建议就是在第3章谈到的5个方面,一般准备充分之后不会出现太大的问题。首先,要想清楚表格所要说明的内容。如果要说明一些自变量对因变量的影响,就要想好行文中要强调哪些变量。例如,假设要研究对于他人的自我暴露和非自我暴露,那么在进一步假设中,需要将性别作为第二个自变量来考虑,这就很可能需要用2×2的方差分析来考察数据。像这样一个简单的四组比较,可以通过文字的形式来呈现数据,而无需使用表格。如果用到表格,则很可能会像表11.1那样来呈现数据,或者像表11.2那样安排准备好的数据。那么这两种表格中的哪一种能更好地表达出模型的重点呢?大多数的读者认为表11.2更好,因为在大多数情况下,左右对比要比上下对比容易。这是一个简洁明了的例子。还可以通过增加额外变量将它变得复杂一点儿,例如,可以区分自我暴露的类型。可能还会有性与亲密、金钱以及专业理想等方面的相关内容。

表 11.2

暴露和非暴露模式下不同性别的自我暴露水平：平均数和标准差

	暴露模式		非暴露模式	
性别	*M*	*SD*	*M*	*SD*
女性	7.60	2.34	4.76	1.89
男性	4.85	2.02	2.67	1.93

注：$n=22$ 人/组

怎样才能够在不同性别及不同模型下呈现这 3 种类型的对应方式呢？表11.3给出了一种可能的解决方案。

由此看来，很容易就可以做到简洁明了地呈现主要自变量的作用。的确，这两种模式的条件是并列的。但是如何比较自我暴露和非自我暴露的两种条件呢？一种是通过测量进行比较，这需要先在左边找到一项，然后看与它对应的右侧位置，而忽略中间的项目。用手指指读表格也许会比较容易，但是采用另一种不同的方式来编排表格也许会更便于读者理解主要变量。

表 11.3

暴露和非暴露模式下男性和女性自我暴露水平的 3 种内容平均数和标准差

	暴露模式						非暴露模式					
	性/亲密行为		经济情况		职业追求		性/亲密行为		经济情况		职业追求	
性别	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
女性	5.20	2.01	4.80	1.98	4.75	1.87	3.26	1.67	3.91	1.56	3.84	1.73
男性	3.35	1.77	5.92	1.96	5.87	2.15	2.67	1.78	5.13	2.21	5.46	2.04

注：$n=25$ 人/组

看表 11.4 中的内容。采用这种表格排版能够使读者比较容易地理解自我暴露效应的模式，因为可以通过纵向对比来比较在每个因变量上的性别差异。此外，很容易通过对比每列的平均值来对暴露模式与非暴露模式两种条件进行总体比较。

一般建议在学位论文中多使用表格。可以将解释说明研究核心内容的表格

呈现在结果部分,而将不那么重要的表格放到附录中。《APA 写作手册》要求在表格的结尾部分标明脚注。在论文排版时,应将表格放置于与相关文字内容较接近的地方。一些学校对毕业论文中表格安排的要求较为灵活。但一般认为,像期刊要求的那样尽量将表格置于讨论这一表格的文字内容附近更为合适。这将有利于阅读,因为这样就不需要前后翻页或者在正文与附录之间来回跳跃以寻找相关表格。

表 11.4

暴露和非暴露模式下自我暴露的 3 种内容:平均数和标准差

	暴露模式		非暴露模式	
性别	*M*	*SD*	*M*	*SD*
女性				
性/亲密行为	5.20	(2.01)	3.26	(1.67)
经济情况	4.80	(1.98)	3.91	(1.56)
职业追求	4.75	(1.87)	3.84	(1.73)
男性				
性/亲密行为	3.35	(1.77)	2.67	(1.78)
经济情况	5.92	(1.96)	5.13	(2.21)
职业追求	5.87	(2.15)	5.46	(2.04)

注:$n=25$ 人/组

编辑表格时,要留意其可读性。在表格中要留出足够的空间,每列的小数点对齐,如果可以,尽量将一个表格呈现在一页上。呈现研究结果的图表不同于整页的图表,整页的图表有时候必须通过连续几页才能呈现清楚,呈现研究结果的图表则应尽量位于一页上。有时候需要用表格来呈现数据的全部细节,那么就不得不分页呈现。在行为科学研究中,几乎不需要报告小数点右面两位以后的数字。省略小数点右两位后的数字可以省空间,并可避免出现不科学的行文内容。如果要整页呈现图表,请务必在页边的标题栏注明图表的名称(比如在左边)。

至于图表的标题,要既简短又能说明图表的内容。不要包含表格名称本身已经含有的信息。对比表 11.3 和表 11.4,这些表格都包含相同信息,只是以不同的方式排列。表 11.3 标题的不足在于包含了过多的细节。这是由于表格名称和标题二者出现了重复信息所致。注意,表 11.4 的标题因为不提及其他自变量(性别)而显得简洁明确,由于表格名称已经包含了这些信息,因此不会造成任何困惑。而在用词上,像"水平"和"分数"这样的词一般都可以省略,因为读者很容易就能根据表格内容对其进行推断。注意,图表中一般倾向于将因变量放在首位,其次才是其他主要

统计数据。因此,像"在暴露模式和非暴露模式下自我暴露的性别差异:平均数与标准差"这种说法要好于将因变量和统计数据混合在一起呈现。这就是"主旨第一,统计第二"的准则。

如前所述,呈现有效的表格需要练习。《APA 写作手册》提供了许多关于表格的详细介绍,甚至还包括一个具体的清单表(见 APA,2001,pp. 175-176)。在开始着手编辑表格前仔细研读这一宝贵资源将会很有帮助。贝勒和莫斯特勒(Bailer & Mosteller,1988)及尼科尔和佩曼(Nicole & Pexman,1999)为各种分析提供了良好的示例表格。模仿专业杂志上其他研究者进行相似数据分析时使用的表格也会有所帮助。切记要选择最清晰的表格模式。另外,也可以多尝试不同的表格排版形式,然后再请他人评阅,并对表格是否易于理解提供反馈。

一旦决定采用某种特定的表格排版方式,那么最好用完全相同的方式来呈现所有相似的数据。还记得以"简单地重复"的方式来报告研究结果的建议吗?这一原则同样适用于表格。不要让自己和读者在表格的格式上花费太多的精力,一个精心准备的表格可以作为范本在很多地方应用。另外还要记住:论文中的表格只呈现最重要的数据,如个别被试的数据和其他详细资料等一般都在附录里呈现。

准备合适的图表

呈现统计数据的其他方式主要是图表。除了文字和表格,论文中还可以呈现图形、图表、图片和图纸。因为图表比文字和表格更难于准备,所以最好尽量少使用,只有在难以用除图表以外的任何其他方式传达某些信息时,才使用图表。例如,一个学生的毕业论文研究餐桌教育的作用,论文中呈现的、认真准备的图表能清晰地表明餐桌教育作用的分布情况(Nelson,Cone & Hanson,1975)。用图表来描绘的自变量之间的交互作用和随时间推移的趋势十分有效。

在使用图表前问自己下面的问题:

1. 想通过图表来表达什么?
2. 用文字或表格能否阐明这些内容?
3. 图表中的信息是否与其他方式传达的信息有重复?
4. 是否可以将这一图表中的信息放到附录中呈现?
5. 何种类型的图表(曲线图、图画或者图片)最合适?
6. 如果采用曲线图,哪种曲线图(柱状图、条状图还是线形图)最合适?
7. 如何生成图表(如,手工绘制、计算机软件生成或者由专业人士绘制)?

回答过这些问题之后就可以开始准备图表了。和表格一样,制作图表也有一些惯例。制图前最好花几分钟来思考这些要求。本章将不再讨论各种可供选择的图表类型。《APA 写作手册》(2001)以及 Parsonson 和 Baer(1978)、Tufte(2001)、Cleveland(1995)等都对此进行了明确的阐述。这里只结合一些普遍接受的惯例与个人偏好,阐述使用图表时需要注意的几个方面。

首先要做出直角坐标图(两条线呈直角)。纵轴或称 y 轴表示因变量,横轴或称 x 轴表示自变量。一条经验法则是,纵轴的长度大概要是横轴的两倍。纵轴应该包含因

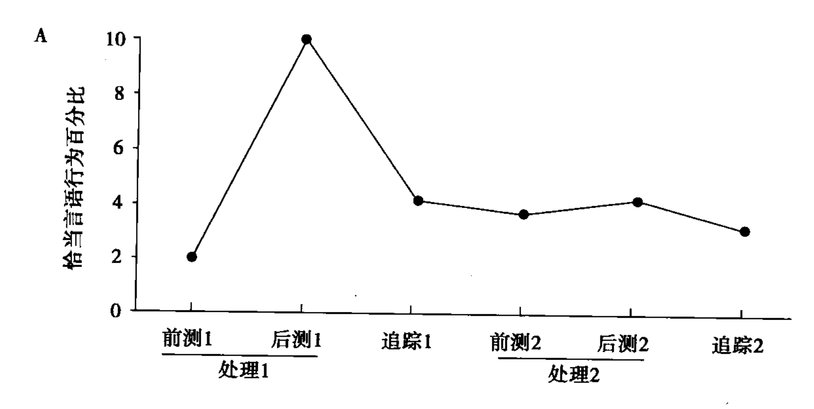
变量的整个范围。举例来说,如果要呈现的是百分比,则坐标轴的范围应该是从 0 ~ 100。如果分数只集中于某一部分,可以考虑只显示这一部分的分值。图11.1说明了这点。

图 11.1A 中的数据与图 11.1B 中的数据会给读者不同的印象,这同纵轴的取值范围有关。图 11.1A 中前两个数据点所显示的前后测差异看起来很大,而当像 11.1B 图一样给出了数据的整个范围时,同样的差异引发的感觉就不同了。

相同单位在坐标轴上的距离应该是相等的。表现在横轴上,就是相同的距离表示相同的时间量。注意图 11.1,前后测 1 和前后测 2 两者在横轴上所占的长度是相同的。而从后测 1 到前测 2 的距离则并不相同,因为这反映了不同时间量。

坐标轴上的单位一般从两轴交点开始,沿箭头方向递增。一些研究者(如, APA,2001;Parsonson & Baer,1978)倾向于剔除落在图外的数据,其他人则倾向于将之纳入到图表中。无论如何,纵轴不应低于最高的数据点。还要为两轴做出清楚的标签,标明变量及变量单位。在图 11.1 中,变量是适当的言语行为,测量的单位是百分比。标签应与坐标轴平行。

图 11.2 说明了关于图表的另外一些规则。再次提醒读者:要注意纵坐标不连续的情况。现在将介绍横坐标不连续的情况。



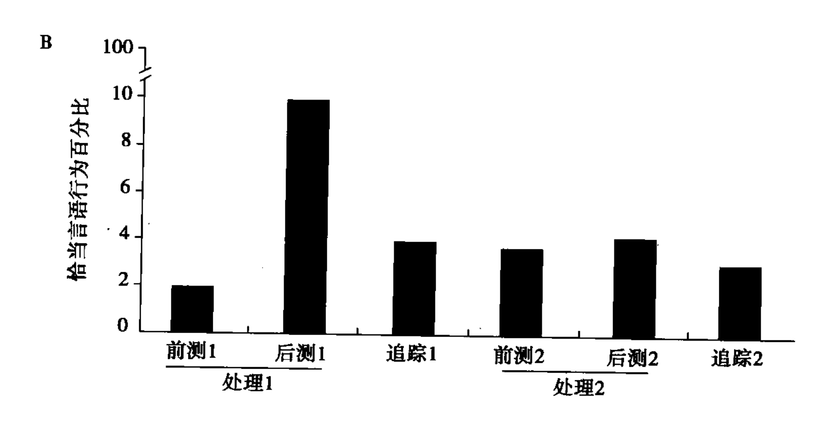


图 11.1

A 用线形图呈现离散的数据,*y* 轴取值连续,无断点。B 用条形图呈现同样的数据,*y* 轴取值不连续,有断点。来自"The Analysis and Presentation of Graphic Data",By B. S. Parsonson and D. M. Baer,1978,*Single Subject Research*(p. 117),San Diego,CA:Academic Press,Copyright 1978,使用经过 Elsevies 同意。

请注意，图 11.2 的总体布局与研究中实验设计的类型（*A*—*B*—*A*—*B*）相关。同时要注意每一段的标签和纵向分割的虚线。对每段的内容进行命名是非常有用的（例如，信息反馈），因为这样可以比一个总体标签传达更多的信息（例如，治疗）。

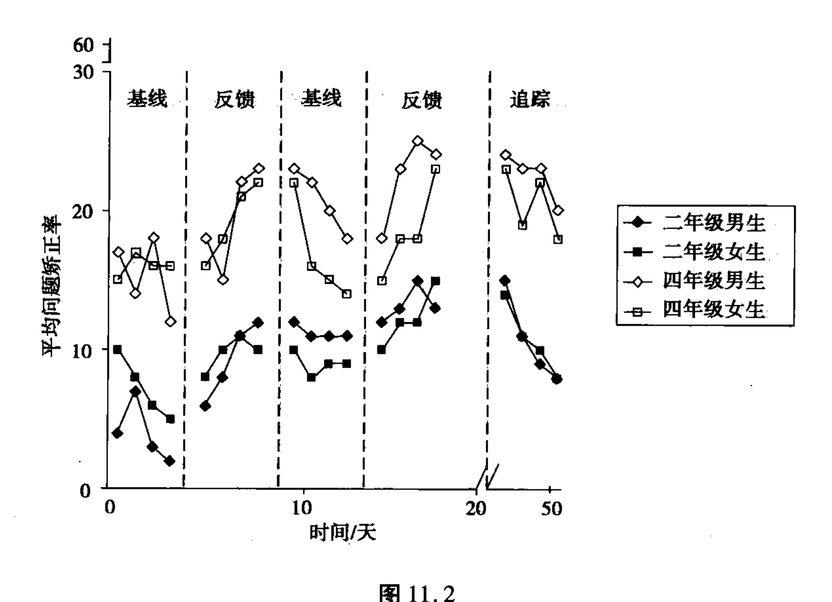


图 11.2

二年级和四年级男女生在 4 个阶段的学习及追踪测试中的平均问题矫正率。

图 11.2 呈现的数据来自于 4 组被试。事实上，可读性较好的图一般最多包含 4 条线图。如果数据接近，4 条线也可能是多余的。注意，这里还包括一个基线组。把基线组的数据纳入到图表中不失为一个好主意。但如果空间不允许这样做，则最好将它呈现在图表的右侧。

一旦决定采用某种类型的图表，那么最好在全文中都一致地采用同种图表。图 11.2 中，方块表示女生，菱形表示男生，而分别用空心和实心符号代表二年级和四年级组，在不同水平上保持使用相同的符号来代表同一变量。如果文章中使用一个以上的图表，要保持符号一致。即，一直用方块代表女生，菱形代表男生，用空心和实心符号代表二年级组和四年级组。

通过图 11.2 也可以说明其他一些有用的图表使用规则。比如，不同区间的数据点不相连，以便于比较不同区间的数据；用空白来表示缺失值，而不应用指向任何一方的缺失值；数轴上线的宽度应小于数轴本身等。有时还不妨在同一图表中呈现多个变量的信息，但是当变量间的测量方式不同时，制作这种图表可能会非常棘手。图 11.3 给出了这一问题的解决办法。在这个图表里，使用两个纵坐标，一列在左边，一列在右边，这样就可以分别显示出两个因变量由于不同的处理而产生的不同变化趋向，这种呈现方式使变量间的对比更容易。对整个论文来讲，用一个图呈现两个变量比用两个不同的图分别呈现这两个变量更为经济实用。

最后，图表的选择应该与数据相适应。事实上，线图不适合呈现不连续的数据，图 11.1 的上半部分说明了这点。正如帕森森和贝尔（Parsonson & Baer，1978）指出的那样，“数据点之间的时间距离无法确定”（p. 117）。在这种情况下，图 11.1 下半部所示的条形图则更为合适。

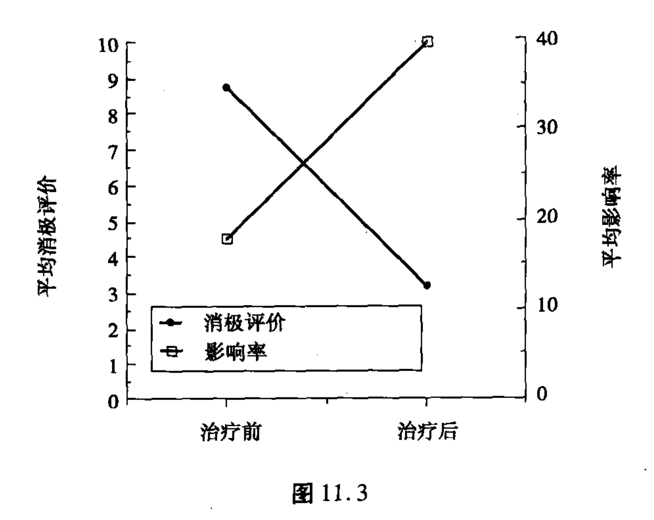


图 11.3

抑郁患者治疗前后的平均消极评价(左侧纵坐标)与影响率(右侧纵坐标)。

对大多数人来说,建立图表的最后一步是对其命名。

在写作投稿论文时,APA 要求将图表的标题单独印在一页上并附在论文的最后,图表本身也一致地放在论文的最后。这样在论文中,空间就可以更宽松。如果将图表放在文中,则一般建议像表格一样呈现在相关文字的附近。图表的题目则呈现在相应图表的下方。

图表题目有两个作用:一是为图表提供了一个标题,二是解释这个图表本身的内容。图表及其标题都是根据其所要呈现的数据而定的。也就是说,读者应该无需在上下文中进行搜索就能理解图表所要传达的信息。标题要描述图表的内容,这通常可以通过简单的句子或短语来表达。像《APA 写作手册》要求的那样,可以将说明性信息放在括号里,呈现在图表结尾处,或者策略性地呈现在图表中。举一个标题过于简单的例子:"图 11.3,言语行为与影响水平对前后测的影响",可以将这个标题与其他图表的题目进行对比。

绘制图表也需要反复实践。《APA 写作手册》谈到了许多与图表相关的细节。还包括专门的图表使用清单(见 APA,2001,p.201)。

与表格一样,进行各种尝试之后再确定图表的最终形式是个好主意。可以请朋友或导师提供阅读反馈。然后修改图表,再重新画图,最后通过软件或者请专业人员绘制最终版本。通过计算机程序绘图是一种便捷有效的方式,大多数主流统计软件包(如,SPSS,SAS)和一些电子统计表格程序(如,Excel,Calc)都有绘图功能。但是,直接剪切和粘贴统计程序中的图表通常并不合适。必须要先对图表进行编辑,以便使字体、线条粗细等都适合。

电脑及相关绘图软件可以直接绘制出符合论文或者出版要求的图表。建议研究者,尤其是低年级的研究生,花一些时间学习使用计算机来绘制图表。这是很值得为之付出努力的,既可以节省时间又可以节省请专业人员制图的费用。此外,最好将自己绘制成功的图表张贴在书桌周围,以显示练习的效果和对程序的掌握,鼓励自己。还要说明的是,上面介绍了主流软件的绘图功能,是因为这些程序的日益完善使之能够绘制出比较全面和复杂的图表。在本章后面会给出一些常用的计算

机制图程序，但这并非此类软件的详尽清单。经验证明，制图程序的更新相当快，并且很容易过时，本书前一版中所列出的软件就有很多已经更新了。获取软件的一个比较好的信息源是与其他同学交流制作图表时所使用的软件，或者可以咨询一下学校计算机中心的工作人员。

如果决定采用老式的手工绘图，那么一些材料可以令这项工作变得轻松。其中，最有用的是蓝线绘图纸和模版。使用时把模版压到纸上，或通过复写纸来作图。这种蓝线绘图纸可以起到不可估量的作用，有助于设计和绘制图表内容，而且当最终复印成品时不会留下蓝线的痕迹。大学里的书店一般都有各种型号的绘图纸及相应工具出售，这些工具包括画坐标轴所用的压线和在线性图上标记所用的符号。耶茨（Yates，1982，pp. 165-171）对手工制图提供了一些有用的建议。

补充资源

书籍和论文

Carr ,J. E. ,&Burkholder,E. O. (1998). Creating single-subject design graphs with Microsoft Excel. *Journal of Applied Behavior Analysis*,31,245-251、

Henry,G. T. (1994). *Graphing data.* Thousand Oaks,CA:Sage.

Heppner,P. P. ,& Heppner,M. J. (2004). *Writing and publishing your thesis,dissertation,and research.* Belmont,CA:Thompson-Brooks/Cole.

Tufte,E. (2001). *The visual display of quantitative information*(2nd ed). Cheshire CT:Graphics Press.

Wallgren,A. ,Wallgren,B. ,Persson,R. ,Jorne,U. ,& Haaland,J. (1996). *Graphing statistics and data.* Thousand Osks,CA:Sage.

Wilknson,L. ,& The Task Force on Statistical Inference. (1999). Statistical methods in psychology journals. *American Psychologist*,54,594-604.

软件

Excel:http://www. ms. com(注:Excel 是微软办公软件套装程序中的电子表格应用程序。它有绘制图表的功能，它包括 OpenOffice 电子表格所要求的、包括计算在内的全部功能。如果不熟悉图表中的某一部分，则建议先探索 OpenOffice 的用法。它是免费的，而且完全符合微软办公软件的格式和要求。)

OpenOffice:http://www. openoffice. org/

SINGWIN:分析单因素设计数据的程序，包括统计和图表两部分;见 Bloom,M. ,Fischer,J. ,& Orme,J. E. (2005). *Evaluating practice:Guidelines for the accountable professional*(5th ed.). Boston:Allyn & Bacon.

思考与练习

呈现研究结果

- □ 报告与研究目的相关的数据
- □ 以逻辑有序的方式报告研究结果
  - —结果的排列顺序
    - · 报告预分析
    - · 报告主要分析
    - · 报告补充分析
  - —呈现相关信息
    - · 统计的名称
    - · 统计相关细节
    - · 显著效应的统计值
    - · 平均值
    - · 标准差
    - · 样本量
    - · 置信区间
    - · 效应量
  - —阐明研究结果
  - —根据心理学论文呈现数据的要求制作出设计良好、清晰的表格
- □ 准备精心制作的清晰图表

12

讨论研究结果

在完成数据分析之后，还要接着前面已撰写的部分继续进行论文的写作，这时可以采取的一个较好方法是先回顾前面完成的文献综述和方法部分，然后再开始论文结果和讨论部分的写作。

回顾文献综述与方法部分

如果在准备阶段已经对文献进行了高质量的、全面的综述,并及时收集了研究数据,则应该不需要进行太大的改动。否则,就需要查找本领域的主要期刊,阅读近期出版的和已经交付出版的论文,以更新参考文献。此外,研究小组在进行讨论时也可能会提出一些需要改进的问题。如果前期的文献综述比较简单潦草,此时则需要付出努力来进一步完善综述部分。

如果在准备阶段撰写的文献综述和方法部分较为完善,后期就只需要进行少量的修改工作。完善方法部分可以从以下 4 个方面着手,使之更加详细明确:第一,如果原来的方法部分是用将来时态写作的,现在要改为过去时态;第二,如果实际操作与研究计划之间存在差异,就要根据实际情况修正关于被试、程序等的描述;第三,添加真实被试的人口学信息(如,通过表格或文字描述每组被试的特征);第四,进行能证明研究测量方法适当性的分析(如,研究数据的内部一致性)。即使方法部分的这几方面无需较多修正,也要回顾文献综述部分。回顾综述部分可以提示研究者应该如何将自己的研究与文献资料相整合。在讨论研究结果的时候,回顾前面曾进行的研究工作尤为重要。

讨论部分的写作

讨论部分通常是写作论文时的最后一项重点内容。学生们有时会觉得不知道应该在讨论部分呈现什么内容——因为似乎已经一切具备了。这种观点不难理解,但是事实却并非如此。讨论部分需要解释研究结果,并将研究结果与研究假设及文献综述联系起来,着重探讨结果的意义和局限。为了使这一过程清楚明确,本章会给出许多关于撰写研究论文中讨论部分的建议,阅读时候要注意思考,记录有用的想法,并考虑如何将这些要点整合到讨论部分中去。

总结研究发现

讨论部分大多是以对结果的概括性总结开始的,但不必重复统计数字,不要陈述 F 值和 p 值,也不要重申分析的技术性细节。这部分所需要做的就是清晰地描述研究结果,尽可能少地使用统计学术语。比如可以这样说,“提示和回忆变量存在显著的主效应。”但是要首先对非专业人士进行访谈,思考怎样向这些本领域之外的、受过教育的人描述研究结果,以使其能够理解所描述的内容?“在学习后参加相关记忆测验的学生会比只学习而没有参加记忆测验的学生记得更准确,回忆成绩更高”这样的描述可以让非专业人员也能够理解。

结果总结的一种较好的组织方法是围绕假设或研究问题展开讨论,这样的总结逻辑更清晰。可以按照前面文献综述部分中排列研究问题与假设的顺序来组织讨论部分的结构。最好在讨论部分复述每个假设,这样读者不必返回之前的段落再看一遍。这种组织结构特别适合于各个假设或是各组假设在概念上相分离的情况。

要注意的是，需小心避免对不同假设的讨论内容出现重复叙述。

结果总结的另一种有效组织方式是将主要的研究发现归纳到一起。这种方式较具有挑战性，但如果能够把握好则会产生不错的效果。要事先问问自己应该如何根据变量和测量组织研究的结果部分，确定好组织方式之后再开始写作。比如，劳拉（Laura Goyer，2005）让两组犯罪水平不同的男孩看其他男孩谈话的录影带，再对比观看高犯罪水平录像带和正常水平录像带后男孩们的反应。研究假设是观看不同类型录像带后两组男孩在自我报告及他人观察上都会出现差异。研究者收集并补充了观察资料和纸笔记录数据资料，同时观察观看录影带时男孩的注意情况及交互影响。

但劳拉的研究结果并不支持她最初的假设：不论是假设预期的变量上还是补充变量上，看不同录像带后两组男孩的反应都没有差异。相反，几乎所有变量的变化都取决于录影带中的谈话类型，有些是违规性的，有些是一般性的。因此，根据这两项研究发现来组织讨论就会更有意义。

像劳拉的研究一样，有时研究结果并不符合预期——这时就需要进行计划之外的分析或者需要讨论出乎意料的研究发现。有时候，恰恰是不符合预期的内容提供了更有趣的发现。如果对被试进行组间比较，是否会发现组间存在非预期的差异？这些差异与因变量有关吗？比如，一位同学（Sikora，1989）调查父母离异及近期生活改变与学龄前儿童同伴行为之间的关系。结果却发现，父母离异的孩子和生活上出现重大改变的儿童更多来自较为贫穷的家庭。令人惊讶的是，离异及生活改变与儿童行为的关联很小，但与社会地位却有显著关系。因此，尽管研究假设根本没有涉及社会地位，接下来却要重点对其进行讨论。

解释研究发现

概括研究结果仅仅是整个讨论的一个组成部分。概括之后还要说明这些研究结果意味着什么？体现了自变量和因变量之间怎样的关系？是否存在相关？这种关系是只适用于一个因变量还是可以用于描述多个变量？这些研究结果是否存在混淆？有些研究能够确切地证明预期，研究发现与预期十分一致。在研究设计阶段就应该仔细思考这些问题，以使后面的工作更为容易。走运的话，现在就可以进行到为研究发现选择理论解释的阶段，并探讨这一研究对于理论、实践和整个世界的意义。

但是，大多数研究并不那么幸运。有些部分可能会如愿，另一些部分则可能不符合预期。不符合预期的部分也有其价值，可以提供许多供讨论的内容。在这种情况下应该考虑哪种变量能够支持预期的关系，哪种变量不能，以及出现这些情况的原因。显著的变量之间有哪些共同特征？要指出这些共同点。为什么有些变量不显著？要把显著和不显著的结果相比较并区分开，仔细思考出现差异的原因，或许就会有所发现。

可以从以下几方面考虑测量结果与研究预期不完全一致的原因，并做出解释：①工具问题（选用的工具是否能够很好地测量因变量？所采取的测量是否能可靠地反映真实的效应？）；②自变量缺乏文献支持或缺乏控制（是否对自变量进行了良

好的筛选和控制？实际操作与计划的控制是否一致？自变量足以引起这些差异吗?)；③样本的规模(研究样本量是否合适，是否足以揭示某种效应?)；④研究所采用的特殊程序也许会掩盖自变量的作用；⑤自变量与因变量之间的真正关系可能不同于预期(是否需要重新考虑自变量所起的作用？是否需要另外的调节和控制?)。依次仔细考虑上述每一个问题，讨论与预期不符的内容及出现这种情况的可能原因。如果研究者在之前的收集和分析数据时坚持记录了研究日志，此时可以回头阅读这些记录，看是否能够获得启发。在考虑做出某一解释时，还要想一想应如何用自己的研究结果去支持或者否定它们。对于不符合预期的情况还可以进行补充分析。例如，如果在研究中应用了相对较新的测量方法，信效度信息不充分，则可以将这种测量方法与其他信效度较稳定的测量方法相比较。二者之间的关系也许可以支持或者否定这种新的测量方法。

有时候，经过几个月的辛苦工作，仍然不能得到显著的研究结果，那就要找出这种情况的原因，此时可以考虑上述因素。比如，对数据进行可行的补充分析也许会得到有用的发现。例如，当研究数据有较大变化时，研究者可能会怀疑只有一组被试的反应符合预期，那么就可以单独分析这一组的数据；如果怀疑是某种人口统计学变量对结果产生了影响，则可以计算这一变量与因变量之间的相关。在结果部分也要报告这些补充分析，并进行补充讨论(如果样本量太小则可以考虑增加被试)。

如果认真思考之后还是找不到结果不显著的原因，也不要失望。科学是在信息积累的过程中前进的，这其中也包括不符合预期的研究所提供的信息。一项样本规模适当、变量可靠有效的研究设计即使没能得到预期的结果，也还是与符合预期的研究一样重要，可以有所贡献。在这种情况下，不合预期的结果可能暗示着研究假设需要修正。但是要注意，非显著性的研究结果并不一定支持相反的假设，还要考虑到数据的趋势及程度。

在一些研究领域中没有得到显著性的结果时，可以考虑分析人口学变量与因变量之间以及各因变量之间的相关。这可以为讨论某些研究发现及解释不符合预期的结果提供更多信息。

在论文中呈现研究发现

任何研究都是在大量前人研究的基础之上进行的，写学术报告时可能需要以相当大的篇幅来论述相关研究。因此，不要将论文的讨论部分写得让人误以为这是该领域中的唯一研究。当总结论述研究结论时，要认真考虑它们与先前的文献综述是否相符。可以从被试特征、测量工具、自变量(如何控制自变量)、研究设计和实验程序等方面着手，将自己的研究与先前的研究相比较。然后认真考虑这些研究发现与自己的研究结论是否一致，这样做有助于澄清现有研究中的矛盾或者与过去的研究结果相冲突的结论。

如果你的研究结果与前人的研究相一致，那么与前人研究在方法上的不同之处也许会扩展先前的发现。例如，已经有研究者在实验室中通过重复测量的方法发现夫妻归因与婚姻满意程度相关。这时如果通过归因日志记录法进行研究也发现相似的结果，就可以讨论在自然环境里采用这种方法也同样发现了归因与婚姻满意度

的相关。

研究结论有时还有助于澄清现有研究间的矛盾。如果研究有这方面的意义，一定要指出来。当研究者在人口学变量、测量领域和自变量等方面进行更严密的控制时，经常能解释现有研究结果之间的矛盾。如果在研究过程中进行了严密的控制，则研究发现可能会帮助读者更好地理解现有文献。例如，有某些研究采用了男性和女性被试，但不考虑性别间的差异，并得到相互矛盾的结果。这时就可以仔细探索性别差异的影响，可能会得到显著差异。这有两层重要含义：①未来的研究应该把性别看成重要的变量；②过去的研究也许因为混淆了性别变量而没有发现应有的效应，因为这一效应可归因于性别。如果过去的被试样本以男性居多，则可能是男性的特征在其中起了主要作用；而女性被试居多的研究则可能更多地反映了女性的特征。

如果自己的研究发现与其他研究不一致，就要仔细思考其中的原因。你的研究和别人的研究有什么不同？可以考虑方法学方面的解释：不同的测量方式、步骤或者样本能够说明差异的原因吗？讨论一下：不一致的研究结果是否能够说明先前研究的推论存在局限？是否说明以后的研究要进行某种控制或辨别？如果这种不一致不是由于研究程序问题或者测量问题等人为因素造成的，那么这项研究也许发现了关于这一研究问题的重要信息。

另外，还要好好考虑研究结果对现有文献的贡献。一篇论文应该能够提供新的学术信息，即使只有较少的新发现，也还是会在现有知识的基础上有所发展。不管认为自己的研究存在多少问题，都要相信它会以某种方式做出贡献，写作时一定要指出这一点。本研究与前人研究有何不同？它如何改进了过去的研究？它对这一领域的研究会有怎样的影响？

考虑研究结果的意义

研究发现有什么意义？它们怎样提高了对相应研究领域的理解？怎样改进了探究假设的方法？当考虑研究发现的深层意义时，要考虑其对理论、研究和实践的贡献。

在理论方面，首先要考虑本研究领域中的一般理论模型。这些模型的基本假设是什么？有什么潜在的前提假设？这些理论对研究发现有怎样的预期？研究结论是否和其中的一两个理论相符？为什么相符？为什么不符？研究发现通常只能支持或者不支持某种理论，而不能证实或者否定该理论。如果研究发现只支持一种理论，这种情况很好。但如果可以同时用几个理论来解释也许同样很好——可以据此对研究发现进行讨论，并展望将来的研究者可以如何设计研究方案。

在研究方法方面，实验控制对深入理解研究发现有很大的作用。你的论文是否提出了实验中需要控制的重要因素或者需要进行辨别的一些因素？比如说，斯科拉（Sikora，1989）早期关于幼儿园儿童的学位论文，发现社会地位与社会行为的相关有两方面的含义：①未来的同伴研究应该控制社会经济地位；②应该明确描述被试的特征。

在研究的实验设计和测量方面，要考虑自己的研究发现对这些领域是否有所推

进?如果你的研究中采用了观察法,而其他研究者采用的都是自我报告法,那么研究结论是否能表明未来的研究应该采用观察法?研究通常都是从一般到特殊,从不完善的实验设计到精巧的实验设计这种不断进步发展的过程。为了更好地理解这一点,可以假设一项采用纵向实验设计的研究结果不同于采用交叉设计的研究,那么就可以认为采用较完善实验设计的研究得到的结果更具说服力。

研究结果有什么实践意义呢?心理学领域的研究结果会对临床、教育、工业或者组织等方面的心理学实践产生什么样的影响呢?什么样的人会对研究结果感兴趣,为什么感兴趣?这一研究的结果如何对潜在读者产生影响?即使研究结果和预期不一致,也还是能说明一些问题。比如,研究结论可能会对一般人的习惯性思维提出挑战,如在关于酒精中毒者子女的行为和心理问题研究时,或者在对被抑制记忆的真实性和准确性的研究时,可能得到不同于日常观点的结论。除此之外,研究者在研究假设的基础上进行的干涉研究有什么意义?对于非本专业的读者来说,一本心理学通俗读物上所介绍的研究意味着什么?

一些研究可能会比另一些研究更有实践意义。例如,临床研究或者商业性研究通常应用价值较强;而生理心理学的基础性研究也许没有什么实践意义。如果研究发现对实践有着潜在的启示,则一定要考虑其应用时的局限。研究者也许希望将研究结论推广到广泛的实践应用中去,但同时要在论文中说明推广的前提条件,并说明在实践中应用这一研究结论时还需要做哪些工作。

在讨论环节中,研究结果的推广是个重要的问题。许多学生在讨论中过于拘泥于研究数据,以至于讨论部分的内容很少,甚至会觉得没有什么值得讨论的内容。在讨论研究发现时,要平衡科学质疑与结果推广之间的关系。实际上,只要不脱离实际研究数据,就可以大胆地推广研究结论。另外还要注意推广研究结论时所用到的假设,并在论文中指明这些假设;要尽可能通过自己或者其他人的研究结果来证明自己的假设。这部分的关键是要清楚地说明研究的推论,并且对支持或者反对这一推论的证据进行讨论。

指出研究局限性

讨论部分是巧妙地夸赞自己研究设计的优越性及研究结果的重要性的地方。但与此同时要认识到任何研究都不是完美的。要写一些比较"谦虚"的内容,指出自己研究的局限。这有两点益处:①能够证明研究者理解研究方法学和研究讨论中不可避免的问题;②也许会使研究者在答辩之前或编辑审稿之前发现自己研究中存在的问题。

一般可以从两方面去考虑研究的局限:①研究过程中的参数选择;②研究实践中遇到的问题。在参数选择方面包括研究程序、测量和研究设计等,所设置的参数在具有研究意义的同时也常常会带来局限。在讨论部分,要仔细权衡参数设置。如果研究过程中出现研究者曾力图避免的、潜在的问题,也可以谦虚地指出自己是如何处理或应对这些潜在局限的。

在收集数据时,也可能会出现一些限制研究结果有效性的问题。比如研究的实际样本量可能小于预计样本量,这会影响统计力。愿意参加研究的被试也许很少,

而且一些被试可能会中途退出；不同被试组之间可能存在人口学上的差异；研究中也许忽视了其他一些需要注意到的混淆因素；自然环境中的观察法也许不可信，等等。这些确实都是研究中令人头痛的问题，但可以将之呈现在讨论研究局限的部分。注意，这部分的关键在于“讨论”，而不是抱怨。对研究中遇到的问题的抱怨不会激发读者兴趣，而讨论这些问题是否影响研究结果，如何影响研究结果，也许更能够引起读者的兴趣。

在研究中可能存在4个方面的限制性因素。第一个因素是研究的设计，这会影响到研究的内部有效性问题，关系到是否可以将研究发现归因于自变量的变化，是否有潜在的混淆变量在起作用。第二个因素则是外部有效性的问题，即研究结果在其他群体中的推广，这涉及被试任务、研究情境、自变量的水平、非实验性影响等。第三个因素包含信度、效度和研究测量等方面，以及一些测量中可能会对研究结果产生影响的任务设置或其他问题。第四个因素则涉及研究所采用的统计分析方法。如果在这些方面都没有限制性因素，还要考虑研究的优点和不足。

注意内部效度限制

研究设计经常采用折中方法。任何相关的、准实验的和不涉及随机分配自变量的设计都不能说是由自变量变化而引起的因变量变化。除非对此进行了控制，否则不能说明自变量对因变量有任何的影响。如果研究采用非实验设计，则必须考虑对自变量和因变量之间关系的解释。包括反因果关系（因变量引起自变量的变化），第三变量因果关系（其他变量同时引起自变量和因变量的变化），相互的或循环的因果关系（自变量的变化引起因变量变化，同时因变量的变化也会引起自变量的变化）。如果控制了潜在的变量，就可以排除一些可能的解释，给因果关系假设以强有力的支持。但仍不能根据这种设计的结果得出明确的因果关系结论。

在混合设计中也存在因果关系的归因问题。这种实验设计涉及至少一个实验性自变量（即研究控制的自变量）和另一个非实验性的变量。例如，假设一项要求老年被试和年轻被试参加两种实验性记忆测试的研究。两种记忆测试中的不同表现可以归因为实验条件的不同。但是与年龄相关的差异却不可归因为年龄因素，因为实验中没有对年龄这个变量进行控制。这方面内容可以查阅卡斯丁，沙迪斯，库克和坎贝尔等（Kazdin，2003；Shadish，Cook & Campbell，2002）所著的关于真实验设计和准实验设计的书，来考察特定研究设计的内部效度问题。

回顾你的研究，也许会发现缺少帮助理解研究结果的信息，或者存在着没有控制好的潜在混淆变量。如果论文前面没有进行说明，则适于在此进行讨论。

注意外部效度限制

即使最仔细控制的实验设计也不会没有局限。真实验设计所需的严格控制常常会影响研究的外部效度，包括环境、刺激、程序和真实生活情境等都是可以进行讨论的适宜议题。

除了设计，研究中还可能存在其他一些影响因素限制研究结果的可推广性。这包括被试的人口学特征和自变量控制等因素（研究结果是否可以推广到其他控制方法或变量中去？）。由于生态效度问题，模拟场景中和实验室中的研究结果的可

推广性存在着局限。

注意测量问题

像研究可信性和被试人口学信息可靠性一样，测量也可能会存在局限。一个普遍存在的问题是，研究者可能会在缺乏数据支持时使用自我报告法替代客观测量来研究可观察的行为。如果研究中存在潜在问题，则需要考虑测量过程、信效度、数据缺失及研究偏向等可能影响研究结果的方面（如，社会倾慕研究中的被试反应设置）。最后，还要考虑测量方法是否能够很好地测量变量：所选仪器设备是否适合于实现研究的目的（有的研究者可能会选用文献中广泛使用但并非最适合自己研究的设备）？这时候，就要考虑这些不足会如何影响研究结果。

注意统计问题

数据分析部分也可能存在问题。收集到的数据是否符合统计假设？统计是否过度保守或者过度自由？样本量是否限制了除最强效应之外的统计检验力？是否通过限制进行统计检验的次数，或者通过设置 α 值的水平等方法控制了 I 型错误？是否使用了新的但是缺少前提说明及详尽细节的统计程序？最理想的情况是不要遇到这些统计问题。但如果遇到了问题，就要回顾论文的结果部分并在向答辩委员会提交论文之前解决它们。如果最终仍存在较小的统计问题，就放在论文的讨论部分进行讨论。

在讨论部分探讨研究的局限也有其危险性，因为研究者可能会认为自己的研究存在过多问题而没有价值。这种想法是不对的，没有哪个研究是完美的，每个研究的结果都会有其局限性。要获得一个确定的结果需要大量的研究证据，单独的一个研究是办不到的。但还是要尽量完善自己的研究设计和实验控制。此外，换个角度想，不足之处也有其积极的方面。例如，严密控制的实验室实验在实验控制严密的同时损失了研究的外部效度，而自然环境中的研究和准实验研究在得到较高的可推广性的同时却损失了研究的内部效度。进行研究要在两个极端之间保持平衡，在承认研究局限性的同时考虑所选择的研究设计的优点。

对未来趋势的评论

论文的讨论部分通常以展望未来研究趋势来收尾。这一部分将要介绍根据论文报告的研究结果指出未来研究趋势的相关问题。

研究者经常说一个研究引发的问题多于其能够回答的问题。在撰写论文时要考虑到研究引发的问题，指出将来可以进行的研究，并在讨论部分做详细说明。

多数论文作者认为把研究结果推广到其他人群很容易，可以通过增加研究问题来超越已有研究。在理想的情况下，首先要介绍研究的基础理论。思考需要对哪些附加的自变量进行进一步的研究？需要对哪些自变量的水平进行更深层的探索？需要考察哪些重要的因变量，评估的最好方法是什么？研究设计上有哪些方面需要进行改进？研究者在未来的研究中需要从哪些方面完善研究方法？

另外，对人口学特征、自变量、因变量还需要进行哪些界定？这些界定对未来研

究有什么指导意义？例如，研究结果可能指出某种认知行为疗法会以特殊的方式对不同的想法产生影响。那么在逻辑上，下一步就应该寻找具有较高信度效度的方法来测量不同类别的想法，然后通过这种认知行为疗法进行干预，看该疗法如何影响测量结果。

考虑未来研究趋势时，可以采用的一种较好方法是自问下一步的研究会做什么，并给出答案。想一想，研究结果提出了什么问题？要弄清楚这一问题，还需要知道些什么？如果没有资源限制，还可以做什么？这是对研究进行畅想的地方。然后把想法整合到现实中，并说明研究展望对理论、研究或实践的重要性。

关于论文组织和写作的建议

在确定论文中将要阐述的要点之后，还要将这些想法组织成有说服力的文字。这部分的撰写通常要比综述部分容易，享受这一过程吧！在论文写作的过程中，可能会学到许多关于科学写作的知识。

然而，虽然写作过程很轻松，也还是不能掉以轻心，要记住避免那些讨论部分经常出现的问题。如果能事先用一点儿时间设计好文章的逻辑结构，可能会有助于论文的写作。

对研究作出评价

论文要为自己的研究结论辩护。批评家会对论文中的观点做出怎样的回应？最好事先预测反对的观点并将之纳入论文。这样既可以充实文章的内容还可以证明研究者的思维敏捷。当然，前提是研究的逻辑要清晰。

论文中要明确给出研究的逻辑步骤。报告时不要只是简单地罗列研究结果，这样做是很不明智的。如果在实验研究开始之前提出了假设，就要清楚地列出这些假设。和研究结果相比，研究假设可能看起来很简单，但是要考虑到其对结果意义的影响。研究性论文应该引领读者经历整个研究过程。

避免常见的问题

在论文的讨论部分可能会出现一系列的问题，要特别注意避免下列常见错误：

1. *不要像讨论显著的研究结果那样去讨论接近显著或不显著的研究发现。*即使这些发现的趋势与预期相符，但还是没有达到显著性水平。比如，“男生比女生在接近显著的水平上更加活跃，男生更加活跃的原因可能是……”这种说法就是像对待显著的结果那样来对待不显著的研究结果。但是可以考虑是否存在某些影响预期结果出现的因素。像下面的说法一般更容易接受，“男生的活跃水平与女生没有显著差异，然而二者之间差异接近显著，这可能是由样本大小问题引起的。未来的研究可以通过增加样本量来进一步探讨这一问题……”

2. *当讨论相关设计的研究发现时不要使用表示因果关系的语言。*一些作者在描述相关研究时强调假定的原因，在讨论时使用如“影响”和“产生”这类词汇是不妥当的！如果研究不是真实验设计，就应该使用如“相互关联”、“有联系”和“有关系”等词或短语代替暗指原因的词汇。

3. *不要认为统计显著性和效应大小是等价的。* p 值是0.05,0.01 还是0.000 01是与随机水平有关的函数的一部分,但同时受样本量影响。比如,一个容量为100的样本,0.19 的相关将在0.05 的水平上显著,然而这两个变量只共享不到4%的变化。讨论相关关系时,要小心地将讨论聚焦到有重要意义的相关关系上。这也同样适用于平均数的比较——p 值不能说明组间差异的显著性水平。因此要避免说某一差异的"显著性较高",并以此来讨论某种效应的大小。

4. *要避免将对某件事的自我报告或他人报告看成是真实发生的。* 比如,"住院儿童的母亲比非住院儿童的母亲更多地报告她们的孩子出现异常"的说法要好于"住院儿童比非住院儿童更加异常"。这两种报告对研究发现是两种不同的解释:一种是住院儿童实际表现更加异常,另一种是从母亲的观察或报告来看两种儿童存在偏差。

5. *不要在讨论部分呈现前文中没有出现过的分析或者前文中没有报告过的研究结果。* 如果希望补充一项分析来进行讨论,则要首先在结果部分介绍这一分析及其结果。

6. *讨论部分不要仅仅简单地复述研究结果。* 就像前文中反复强调的那样,要超越单纯的总结,去解释、综合、分析和批判研究发现。向读者呈现自己提出研究问题的逻辑及通过研究解决这一问题的逻辑。

选择一种恰当的组织方式

可以用很多方法来组织讨论部分,但经验证明,有两种组织结构比较适用。第一种结构和此章的前面几节的小标题有些相似,即从简要解释和总结研究发现开始,再将之与现有文献相整合,然后讨论研究的局限性,最后对未来研究发展做出展望。第二种结构是围绕主要的研究发现来组织讨论,然后整合其他研究的启示、发现以及研究局限,并具体到讨论的每一项内容。专栏12.1 给出了两种讨论方式的提纲。在讨论部分的最后还要做出简明扼要的总结,不要泛泛地总结,而是要将讨论的内容整合到一起,并强调研究要点。就像讨论部分的开头那样,在结尾部分也要对研究做出一个总的评价。最受欢迎的一种结尾方式是强调通过这一研究可以加深对某一领域的总体理解,并提醒读者如何理解本研究对理论、研究或实践的推动作用。然后,论文就基本完成了,这真是件值得庆贺的事情!

提交论文终稿

在完成讨论部分,并将文稿提交给导师后终于可以放松一下了。这段时间刚好可以用来处理最后的细节,以使论文符合答辩的要求。比如:检查参考文献,确保在文中所引用的内容都注明了相对应的参考文献;写论文摘要(要了解本校对论文摘要长度和版式等的要求);修订论文目录(要记得补充图表和表格的目录);准备标题页和签名页等。在答辩之前通常要向答辩委员会提交一份除了致谢页和声明页之外的学位论文完整副本。

专栏 12.1

讨论部分的两种框架

框架一：根据假设组织研究发现

Ⅰ. 结果总结和整合

A. 假设一

B. 假设二

⋮

Z. 其他研究发现

Ⅱ. 对结果的解释

A. 假设一

B. 假设二

⋮

Z. 其他研究发现

Ⅲ. 将研究发现与现有文献相整合

A. 一致的研究发现

1. 介绍研究结果与其他相关研究如何一致
2. 解释为何一致
3. 说明这种一致性说明了什么

B. 不一致的研究发现

1. 介绍研究结果与其他相关研究如何不一致
2. 解释为何不一致
3. 说明这种不一致性说明了什么

C. 研究发现对现有文献的贡献

1. 研究发现如何解决了当前存在的问题
2. 哪些结果支持这些新的差异和控制

Ⅳ. 研究发现的意义

A. 理论意义

B. 研究意义

C. 应用意义

Ⅴ. 研究的局限

A. 研究设计和内部效度

B. 外部效度和可推广性

C. 数据分析与统计力

D. 测量

Ⅵ. 研究方向展望

框架二：根据论题组织研究发现

Ⅰ. 总结和讨论研究发现

专栏 12.1 (续)

A. 研究发现一
1. 总结其与假设的一致性
2. 解释研究发现
3. 与现有文献的一致或不一致
a. 对一致或不一致的解释
b. 一致或不一致的意义
4. 研究发现的局限性
5. 需要特别说明的研究或对研究结果的扩展
⋮
Z. 研究发现 N(参考上述内容)
Ⅱ. 研究结果的一般意义
A. 理论意义
B. 研究意义
C. 应用意义
Ⅲ. 研究的局限性(包括特别的、前文中没有讨论的问题)
A. 研究设计和内部效度
B. 外部效度和可推广性
C. 数据分析与统计力
D. 测量
Ⅳ. 研究方向展望

然后再核对一下本校对毕业论文的细节要求,不要忽略任何相关表格和要求。如此,毕业论文即将大功告成。

补充资源

Heppner, P. P. ,& Heppner, M. J. (2004). *Writing and Publishing your thesis, dissertation, and research.* Belmont, CA: Thompson-Brooks/Cole. [提供讨论部分的写作样例。]

思考与练习

结果讨论

□回顾文献综述
□修改方法部分
—把将来时态改为过去时态
—使报告符合实际工作
—增加图表或其他信息
□总结研究发现

思考与练习

—避免技术性细节
—使用清晰的语言
□解释研究发现
□将研究发现置于上下文中
—考虑研究发现与已有研究相一致还是相冲突
□考虑研究发现的意义
—理论意义
—方法学意义
—应用意义
□纳入研究局限性
—考虑内部一致性问题
—考虑外部一致性问题
—考虑测量问题
—考虑统计问题
□纳入对未来研究方向的展望
□可以使用这些技巧：
—批判性地思考
—避免常见问题
—选择合适的组织方式
□对最终版本进行润饰
—检查表格及其内容
—写摘要
—添加附录部分
—核对参考文献

13

论文开题与答辩

毕业论文一般要在专家委员会上讨论两次:一次是报告和讨论研究计划(开题);另一次是对论文最终结果的讨论(答辩)。开题一般安排在收集数据之前,答辩则通常在完成数据分析和论文撰写之后进行。学校对开题一般没有硬性要求,但对答辩的要求却通常较为严格,博士学位论文的答辩则更为严格。

在开题时,开题委员会成员会与学生讨论其研究设计。因此学生一般要先准备好完整的开题报告(包括文献综述、研究方法、计划进行的数据分析、参考文献、附注等),通过指导教师的审阅之后,再提交到开题委员会。然后在大概两个星期之后,开题委员会将召开大约1~2小时的开题报告会。

开题一般有两种形式:①隐性的问题解决模式;②小型的答辩模式。在问题解决模式下,报告会主要是由开题委员会成员针对学生提交的书面开题报告提出问题和建议,以使研究设计更加清楚和完善。比如,老师们可能会指出研究设计中潜在的问题、提出改进实验程序和测量方法的建议等。这种开题报告会的深层意义是共同合作,开题委员会成员与学生及其导师共同合作有助于做出更好的研究。在小型答辩模式中,开题委员会成员允许学生为开题报告做辩护。这种开题的重点在于由学生阐述开题报告中的各种基本理论,并说明自己的研究设计。这有些类似于论文答辩的模式。无论采用哪种开题形式,开题委员会的成员都是希望通过讨论来完善学生的研究计划!当然,开题报告会也可以同时包含问题解决与模拟答辩两个方面。

通常情况下,论文答辩要比开题报告更正式、更严谨,因为答辩的主要目的在于评估研究生是否具有获得硕士学位或者博士学位的科研能力。这项评估是根据学生提交的书面材料及其在论文答辩时的临场发挥而做出的,所以论文答辩委员会的成员会在答辩过程中问到以下几个问题:你做了什么?发现了什么?这些发现意味着什么?并且探讨论文中的研究与这一领域中其他研究之间的关系。老师们需要寻找证据来证明论文作者做到了以下几点:①理解自己所做的研究及这样做的原因;②能够以科学的视角思考自己的研究;③能够说明自己的研究对其他人的意义。一般而言,对投稿文章的答辩要比对学位论文答辩的要求更高。尽管一些学生能够从容应对、顺利通过,但仍有些学生在答辩时表现得并不理想,甚至有些学生要从开题报告会之后重新做起。因此建议大家要严肃认真地对待和准备论文的开题和答辩。

为开题报告会做准备

院系在安排论文开题和答辩时可能会很灵活。因此,要对这类会议做好充分准备,首先要了解本校的相关规定及要求。专栏 13.1 列出了关于论文开题及答辩时可能出现的常见问题,认真回答这些问题有助于做出更充分的准备。

专栏 13.1

准备开题报告会时需要了解的信息

- 通常在论文研究的哪一阶段召开这种会议?
- 谁负责安排会议程序和会议场地,准备所需的设备?
- 开题委员会成员应该提前多久拿到书面材料?
- 在会议之前应该填写哪些表格? 提前多长时间填写?
- 是否有由非论文开题委员会成员进行的正式审查(如,将论文邮寄出去接受正式的外审或邀请外校专家来参加答辩)? 如果有,需要填哪些表格,该如何准备?
- 会议形式通常是怎样的? 开题委员会成员是一次问完其全部问题,要求学生一起回答;还是轮流提问,逐个回答?
- 由谁主持会议?
- 会议持续多长时间?
- 学生可以带食品进场吗?
- 学生是否需要做正式的报告? 报告多长时间? 要包含什么内容? 需要使用投影仪或者幻灯片吗?
- 开题委员会成员的提问是只涉及论文内容,还是会涉及论文以外的其他方面?
- 由谁记录开题委员会成员提出的论文修改意见?
- 会议流程一般是怎样的? 开题委员会成员会对学生的回答做出反馈吗?
- 与会人数一般是多少? 是否需要每一位开题委员会成员到场?
- 开题委员会成员是否可以通过电话或是视频参加会议? 如果可以,由谁来准备相关设备?
- 除了开题委员会成员和学生本人,其他人是否可以参加会议?
- 会议怎样结束?
- 学生需要在哪些表格上签名? 由谁负责准备这些表格,要到哪里才能找到表格的复印件?
- 会议之后有什么安排? 要怎样根据会议要求来修改论文?
- 修改完论文之后,是由开题委员会还是由导师来进行再次审核?

了解排版格式

不同学校或院系对开题报告的格式及长度可能有不同的要求。有时候会要求学生先概述其将要报告的内容,在这种情况下要总结研究的逻辑及依据,将重点放在研究方案的方法论部分。开题委员会成员随后将根据研究报告提问,这些问题—

般有3种基本形式:①你打算做什么?②为什么做或者不做某事?③关于某一内容要怎么做?这些问题通常聚焦于科研本身,但是开题委员会成员的提问可能也会涉及研究的价值和意义、本研究领域中已有的发现、研究设计、测量和程序等方法学方面的内容;以及一些与数据有关的内容。开题委员会成员通常还会提出一些改进这项研究的建议。

应该由研究生本人而非导师来回答这些问题并对开题委员会成员提出的建议做出回应。当然,对提问的最佳准备是要确切地知道自己在做什么以及为什么这样做,要仔细思考为什么使用研究方案中的方法和理论。在回答提问时应该陈述自己的理由。开题委员会成员对下面这种答案一般是不会满意的,"每组中包含20名被试,这是因为需要尽快完成工作,没有时间再安排更多的被试",因为这暗示着在研究的科学严谨性上投机取巧。而像"根据预实验数据进行的功效分析结果说明效应够强;预分析说明这样的样本大小足以检测这种效应"这样的回答则会博得人心。

论文开题委员会提出的改进建议通常会在一定程度上改变研究的工作量或工作重心。面对这些建议,首先要思考这是否有助于提高研究的质量。如果能够有所提高,就可以采纳这一建议,这可能会带来可观的收益。而如果开题委员会成员的建议涉及改变论文的方法论内容,则可以进一步讨论和商榷。注意不要轻率地接纳或者忽视任何一条建议,在研究遇到困扰的时候这些建议会显得尤为重要。

一定要采纳论文开题委员会的建议吗?答案是否定的,对不充分的要求可以做否定的回应,比如,"这一建议很值得深思,但是它已经超出了本研究的范围。可以考虑在文章的讨论部分加以说明",然后请求将讨论重点转回到论文中的研究上去。另外,尽管应该从科学研究的视角对论文进行充分的讨论,但还是要考虑研究的实践性和可操作性。下面的例子说明具体应该如何进行说明。

> 如果能够将被试取样严格限定在中产阶级的、已婚的、在职的、已经有了一个孩子并且确诊为患有重度抑郁症(但没有其他药物问题或心理问题)的女性,的确是进行这一研究的理想方法,但是本地这种女性的数量不足以满足本研究的需求。因此我准备在不影响研究的情况下扩大被试取样的范围,这是本研究遇到的特殊问题。希望开题委员会能给我一些关于如何扩大被试取样范围的建议。

完美地回答开题委员会成员的提问需要敏捷的思维和良好的表达能力。如果在这些方面存在困难,就需要争取更多的时间来思考。这时可以通过重述问题的方式来争取时间,例如,"那么,您是建议……",然后再评价一下这一建议的利弊,在这些陈述之后,你也许就能够想好该怎样回答了。第二种争取时间的说法比较直接,"请给我几分钟时间考虑",征得同意后坐下仔细思考应该怎样回答。开题委员会成员们可能会安静地坐在一旁等待你的回答,也可能进行讨论。

认真记笔记

在开题报告会上,研究生一般都能够和开题委员会达成修改的共识,因此就应该记录下这些修改意见。在开题报告会开始之前,开题委员会主席通常会指定负责记录的人员。然后在会议结束之前宣读这些修改意见,以确保得到所有人的

同意。

最后，开题委员会成员可能会要求学生回避，以便其进行讨论。最终，他们可能正式通过开题报告（但还是要根据修改意见进行修改）；也可能要求学生在修改之后再次提交开题报告，直到正式通过为止；完全不同意学生通过开题报告的情况很少。不能通过的原因可能是研究中存在着致命的缺陷，也可能是开题委员会认为研究设计及写作方面存在不足，或者是涉及了严重的伦理问题。没能通过的学生通常就要在下一次的开题报告会上重新开题，而且很可能要面对新的开题委员会。

准备答辩

各学校或者院系对论文答辩的具体过程可能有着不同的规矩和传统。但无论怎样，都应该像准备开题报告会一样，事先了解答辩的流程和内容，这有助于做出充分的准备。本章提供了一般性的答辩总体概况，以供参考。

弄清流程

答辩通常以学生的正式报告开始，这将在后面进行具体讨论。由于论文答辩通常是开放性的，其中一些听众可能不了解参加答辩的论文的内容。因此，要像准备向本领域专业人员做报告或在国际性会议上发言那样去准备答辩报告。比如，要严格控制时间，事先要了解报告的时限并进行练习，这样才能确保在规定时间内完成报告。报告中还可以应用一些视听呈现设备，如幻灯片、投影仪或演示软件（如，PowerPoint™或 AutoPilot）等。

当完成研究工作和研究发现的报告以后，答辩委员会成员开始提问。可能会提出关于研究的程序、某项工作的原因、研究发现以及如何得到研究结论等问题；也可能会问及对此的评论和解释；最后，还可能会涉及关于研究思路的问题。提问大多围绕研究中的重点内容展开。专栏 13.2 列出了答辩过程中比较常见的典型问题。在准备答辩时要着眼于重点问题以及研究方法和统计方法的各个细节。

在答辩结束时，答辩委员会通常会将其他人请出房间。然后讨论学生提交的论文书面材料和答辩表现。讨论的时间可能很短也可能很长。长时间的讨论未必意味着会出现麻烦：答辩委员会成员们可能在争论结果或方法中的一个小问题，而非学生在答辩过程中的表现！答辩过程中，答辩委员会可能建议学生对其所提交的书面材料进行某些修改，遇到这种情况不要沮丧，这很常见，并不意味着将要面对失败（除非答辩委员会认为学生没能完成其研究工作，如果这样，则一般不能通过答辩）。在开题报告会时，曾有专人记录开题委员会提出的修改意见，此时可能会进行核对，以确保开题委员会和答辩委员会的意见一致。答辩委员会讨论之后会公布审议结果。学生的书面材料和口头答辩可能同时通过，也可能只通过其中一个，甚至可能两者都不能通过。如果没能全部一次通过，通常还有第二次机会可以补救。如果一次就全部通过，那么就可以开始庆祝了！

专栏 13.2

口头答辩的一般问题

- 你认为自己的研究存在什么问题? 这对研究有何限制? 在将来的研究中应该如何纠正?
- 如果想改进测量方法、实验程序等,应该怎样做?
- 研究结果接近显著,但并未达到显著。应该如何增加结果的显著性?
- 如何确定某表格中 F 检验的自由度? 这一自由度似乎应为……
- 当前的哪种理论或模型可以最好地解释你的研究结果?
- 如果有人认为可以用某种理论解释你的研究结果,你怎么看?
- 如何解释这一研究结果和其他研究结果之间的差异?
- 你所采用的研究方法对本领域中以后的研究有什么启示?
- 本研究对研究现状有何促进?
- 通过这一研究可以增进对本领域的哪些认识?
- 从逻辑上讲,后继的研究应该如何进行(或如何澄清本研究中不能解释清楚的问题)?
- 本研究对心理学的实践有何启示?
- 如果可以重新研究这一问题,且不受任何资源限制,应该如何进行研究?
- 你下一步将如何处理数据?

如何增加在答辩结束后庆祝成功的机会? 首先并且最重要的一点,是要准备好充分的书面材料,即毕业论文。如果提交给答辩委员会的文稿是经过深思熟虑、认真准备的,而且体系完整、结构组织良好(例如,论文内容明确、打印质量很好、目录包括页码、引用标注完整、格式符合美国心理学会[APA]要求、没有拼写错误、图表与文字相匹配等),那么就可能给答辩委员会留下良好的印象,在此基础上进行答辩是事半功倍的。千万不要因为急于完成个人日程表上的工作就将没有准备好的毕业论文提交到答辩委员会,这很可能会适得其反(Yates,1982),不要以为答辩委员会能够体谅你急于完成研究任务的紧迫感。如果答辩委员会认为研究论文准备得不充分,不可能在学生发现自己的文稿存在问题时给予重新准备的机会。第二,要对答辩过程中可能发生的各种情况做出充分的准备,以此来进行规划,并练习处理可能遇到的各种问题。第三,要制订计划来应对可能影响表现的过度紧张。后面将对这些问题做出详细说明。

报告的准备

答辩开始时的报告为整个答辩过程奠定了基调。可以考虑在答辩前花一些时间写一个报告提纲,然后进行练习,直到你能够在给定的时间内做出专业的、正式的、连贯的研究概述。报告的时间一般限制在 10 ~ 20 分钟,但是也有可能要求半个小时或更久。

报告的主要结构通常与 APA 格式的出版要求相似,包括引言、方法、结果和讨论几个部分。对 20 分钟的报告来说,大约用 3 分钟时间对研究进行概括性介绍,用

6分钟来描述所采用的方法,用6分钟介绍研究结果,最后用5分钟介绍讨论比较合适;对45分钟的报告来说,各个部分可能分别约为10、20、10、5分钟。但也不需要太过拘泥于上述4个部分。口头报告不同于论文写作,需要时可以将其中的一些部分整合到一起。例如,可以在报告结果的同时对其进行讨论。方法和结果部分一般会比引言和讨论部分需要更多的时间。

一些演讲者喜欢以一个笑话来导入报告。这样可以打破过于严肃的气氛,表现出亲和力和幽默感。但是如果你平时就不习惯于说笑,可能就不适合采用这种方式。另一个常用的导入策略是感谢答辩委员会并微笑着承认自己的失误,这时最好频繁地、亲切地和答辩委员会成员进行目光交流。

在报告的开头,要告诉听众你为何对这一特定的研究产生兴趣,这部分要求专业并简短。听众不会对琐碎的问题有兴趣,比如在介绍自己关于肥胖症的研究时述说自己或者家庭成员遭遇的体重问题是不恰当的。更好的表达方式是这样的"作为布里马克思博士研究小组中的一员,我在过去3年里对鲸脂进行了研究。最初对这一问题产生兴趣,是想知道为什么有些人比其他人吃得更多……"

然后简要回顾这一领域的相关研究,介绍研究的背景。这可以参考第5章中关于组织文献综述形式的内容,有逻辑地带着读者进行回顾,并巧妙地得出结论,指出自己的研究正是逻辑上应该在现有文献基础上进行的下一步研究。限定的报告时间通常很短,无法详尽介绍研究的脉络。因此,要假设听众对这一领域有一定程度的了解,然后集中介绍与自己的研究关系最紧密的内容。简略地总结引发这一研究的现有文献,并指出主要的结果。要给出现有研究的一般性综合概况,而不是一连串无聊的题目和日期。然后指出本研究的缺陷和为后继研究奠定的基础,以及自己的设计如何能解决或者克服现有研究的不足。介绍时要保持谦逊,每一项研究都存在问题,都不是完美的。不要过分批评前人的研究,自己的研究也许同样会面临这样的命运。如果论文的文献综述部分用表格来总结现有研究,那么可以在条件允许的情况下应用视听设备呈现给听众。

引导听众同意研究的逻辑之后就可以开始陈述所研究的问题和研究假设了。投影仪或幻灯片可能会起到辅助作用。然后是研究的方法部分。按照传统的APA格式,首先要说明研究被试和实验设计,介绍被试招募程序和选择标准,还要提到被试知情同意。其次要描述研究中对自变量和因变量的测量。需要再次强调的是,如果条件允许,可以使用幻灯片或者投影仪等视听设备来呈现一些信息。要报告测量的信度和效度,需要的话,还要说明如何训练评分者或观察员,以及如何保证其持续高效准确地收集数据。简单介绍如何在被试参与研究之前向其介绍研究程序,研究之后向其做出简要说明的。

报告中要明确介绍研究程序。对于简单的设计,可以口头说明。但对于复杂的设计可能就需要借助幻灯片或投影仪来呈现一些资料。

显然在这样短暂的时间里无法详细描述研究方法的每一个细节,那么就可以考虑抛开听众不容易理解或者不那么重要的细节。例如,在一项观察学习研究中,听众无需知道对每种观察类型中词汇的界定。如果不能确定报告中应该包含哪些内容,就先将最重要的内容放到报告中,然后依次填充次重要的内容,直到填满介绍方法部分所占的时间。

下一步是报告研究的结果。建议按照研究假设的顺序来呈现研究结果。如果使用多因变量测量来检验假设,就要先确定因变量的呈现顺序,然后在考察每个假

设时都按这一顺序进行。同样,因为时间有限,不能报告所有数据的细节,只能选择性地报告与假设关系最密切或者说明重大发现的数据。如果有许多变量证明了同样的结果,要进行总结归纳而非逐一报告。

无论从哪方面看,简单的总是最好的。不要在陈述中报告 F 值和 p 值,只需用简洁的语言说明你进行了哪种统计分析及分析结果即可。另一个简化报告的方法是只呈现与假设及主要研究发现相关的结果,而不逐个介绍所进行的全部分析,也不要过多地介绍不显著的结果。同方法部分一样,可以先筛选出最重要的内容,估计介绍这些内容大概需要多长时间,再在时间允许的范围内充实其他内容。与其报告研究中所有的数字来将听众弄晕,不如进行选择性的简单报告。

在报告研究结果的同时或者之后,要讨论研究的意义。同样还是要按照报告假设的顺序来讨论。从介绍研究目的开始,接着对假设进行论述。如果有多个研究假设,可以考虑把相关假设归为一类。要经常提及自己研究与他人研究之间的关系,以及这一研究结果可能对该领域研究产生的影响。

除了要对研究结果的意义进行深入讨论,还需注意态度要谦逊。可以承认一些研究中不可避免的问题,然后展望未来的研究。可以承认研究中存在明显的不足,并指出对当前研究的影响和限制,但要注意不宜过分地自我批评。这是一种安全的方式,因为任何研究都有其局限性。结束报告的一个较好方法是对未来的研究提一两条明确的建议,但如果前面的讨论中已经融入了这些建议,则不必重复。

准备视听材料

很难想象持续 5 分钟的研究报告没有任何形式的视听呈现。视听呈现的目的是以更形象的方式把想表达的信息传达给听众,使听众保持清醒并且产生兴趣。视听呈现的量和形式取决于环境条件和报告长度。在报告中可以平均每 10 分钟借助投影仪、幻灯片,或者其他设备来呈现一次相关内容。

对视听呈现手段的应用有各种各样的指导手册。首先,要选取可用而且可靠的呈现方式。例如在呈现文字材料,包括表格、图表以及一些必要的介绍时,可以考虑使用投影仪。投影仪是比较常用的视听呈现设备,此外幻灯片也非常方便。虽然使用幻灯片可能不像操作电脑那么容易,但还是较为容易操作的,且成本较低,可以帮助听众理解报告的内容。需要补充的是,使用投影仪呈现时室内可以是明亮的,但是在呈现幻灯片时则需要周围环境暗一点。

幻灯片是一种容易给人留下深刻印象的媒介。使用电脑放映幻灯片,甚至比使用投影仪设备还要简便,所需的经费也少。幻灯片往往比投影仪更能充分传达所要表达的内容,使报告更形象。例如,呈现一个男孩坐在实验设备前进行实验的场景或者数据处理的中间过程等,这些都可以使报告更生动形象。通过练习很容易就可以掌握幻灯片的使用方法。建议另外做一份在报告过程中不会用到,但是在后面回答提问过程中可能会用到的备用幻灯片。备用幻灯片的内容可以包括测量方法、变量的频次分布、数据表格等,在报告过程中不需要呈现备用幻灯片,而在回答答辩委员会提问时可能会派上用场。在报告中应用电脑幻灯片很快就会成为专业报告的要求,所以掌握这些方法十分有益①。

① 现在的专业性报告一般都要求配合用幻灯片呈现相关材料。——译者注

提高报告生动性的另一种有效途径是使用音频材料。这也是一种非常灵活的沟通媒介,但同样需要合适的设备。一些电脑报告程序支持在报告过程中融入影音片段,例如,如果把幻灯片呈现的报告和 DVD 或 VCR 视频相结合,这些程序中的索引系统可能允许快进或者快退,以呈现某一特定的视频或者音频片段。如果决定采用 DVD 或录像带,要确保所有观众都能看清并听清所呈现的内容,就像使用幻灯片和投影仪时要求的那样。

如果你是第一次使用电脑做报告,建议遵循军队中的古老格言:保持简洁。墨菲(Murphy)提出的关于研究报告的守则指出,任何地方都有可能出错,特别是与设备相关的部分。如果使用像 Powerpoint™或 AutoPilot 等的系统,要在答辩前多次核查其细节。确保自己知道如何将做好的幻灯片导入报告时将使用的电脑,如何连接电脑,如何使用投影仪。要确保答辩的场所有可用的投影仪,需要注意的是最好准备一套备用设备,以防投影仪在最后关头出现问题。无论准备使用什么设备,都要在开始报告前再检查一下它的情况。要知道一旦出现故障,比如投影仪的灯泡坏了等状况(请相信,这种情况发生过),要到哪里寻找备用设备。可以试着坐在房间里的各个位置检查声音和图片的效果。要和在昏暗中换幻灯片的工作人员,还有发材料的人员事先做好沟通工作。

在开题和答辩会议上,学生们的幻灯片既有优秀的也有糟糕的。糟糕的幻灯片可能存在一些常见的问题:如幻灯片页数过多(每页幻灯片还含有过多的文字),呈现了过多数据,或者字体过小等。这是严重的问题,因为会使听众因为听报告而忽略了幻灯片中呈现的内容,或者为了认真看幻灯片而忽略了报告(二者不能同时被接受理解)。有的报告者在报告过程中只是读幻灯片上的内容,这也是致命的问题。记住,这些视听呈现只是用来协助口头报告的,而不能完全取代口头报告。

所以制作幻灯片时要有选择性。不能把研究的所有内容都放到幻灯片中,只要呈现一些关键词来点明要点就可以了。不要因为论文中已经有现成的或者因为喜欢论文中表格的格式就直接从论文中截取表格。通常,直接在屏幕上呈现这种信息的时候都不太容易被理解。尽量不要使用不必要的幻灯片表格,但如果确实需要就选择性地呈现重点数据,还要注意将表格排列整齐。多用几个包含较少数据的表格要比用一个表格呈现很多数据好。不要呈现不必要的统计细节(如,F 值等)。要确保最后一排的听众也能看清幻灯片上的数字。

制定战略,处理质疑

许多学生害怕在答辩的过程中被提问。如果能够知道答辩委员会可能会问什么(见专栏 13.2),并提前反复思考、想好答案,则可以控制这种恐惧。或者事先想好回答提问的基本策略也会有所帮助。本章中与开题有关的部分介绍了一些这方面的策略,接下来会列出其他相关策略。

首先,要像科学家一样专业地回答问题。即使与答辩委员会成员熟悉,也不能随便开玩笑。正如耶茨(Yates,1982)中肯地指出,即使答辩委员会成员提出的问题不重要且没有必要,也不能表现出生气或者攻击等不良情绪。需要再次强调的是,要将听众当成是国际性专业会议上的听众,并适当地表现自己。

另外,不要着急。同开题时一样,要弄清楚不易理解的提问,确保明白问题问的是什么,全面思考后再做回答。记住,答辩委员会成员可能没有注意到在报告中已

经出现的内容,因此也可能提一些很容易回答的问题。

回答问题时要简明扼要。只回答被问到的问题,而不要回答希望被问到的问题。不要对问题发表过长的演说,特别不要使用不常见的术语或自己不全懂的概念。使用不恰当的术语正说明自己根本不了解自己所谈论的内容。

大多数学生在答辩时都会面临一个或几个错综复杂、难以理解的问题。如果不清楚这个问题问的是什么,就花些时间来弄清楚,可以请提问者来澄清问题,直到真正理解问题的含义。

有时也可能会出现一些没有标准答案的问题。对于这些问题,就需要猜测提问者的意图,尽可能根据已有信息进行揣测。像呈现事实那样来提出猜测,是回答此类问题的较好方式。

当遇到不会的问题时也不要惊慌。如果真的不知道问题的答案,可以说"不知道"或者"不确定",而不要毫无根据地猜测。偶尔对一个问题不清楚不会影响整个答辩(当然不是"被试是谁?"或者"采用什么方法进行测量?"这种基本的问题)。另一方面,愚蠢的猜测传达着两个方面的内容:①你不知道这个问题的答案;②你不承认自己的无知。

有时可能还要面对有人哗众取宠的问题。研究者阿尔比特·法瑞尔(Albert Farrell),指出答辩委员会成员的提问可能有 3 种原因:第一种也是最罕见的一种,是真正想知道问题的答案;第二种也是最常见的一种,是因为想知道学生是否知道问题的答案;第三种原因是,要向答辩委员会的其他成员展示自己知道问题的答案,这就是哗众取宠的提问,这种问题通常很长,或者会对学生的回答做出反馈,这时的反馈并非意味着答辩者的回答不全面。不要试图与哗众取宠的提问者进行辩论,就让他自我展示好了。但可以进行适当的评论,如"这是非常重要的、有趣的、有意义的观点",并继续进行答辩的过程。

很难正确地辨别出答辩委员会成员对答案是否满意。建议学生努力提供好的答案,而不是费尽心机揣测答辩委员会成员的看法。答辩委员会成员可能会在微笑点头的同时并不满意你对问题的回答,而一个板着面孔的提问者却可能很喜欢你的回答。提前观摩本院系中其他同学的答辩有助于了解此类答辩的氛围是严肃的还是轻松的。当一切进展顺利时很可能出现轻松的气氛,答辩的气氛受答辩委员会成员以及交流状况的影响。当学生准备充分且表现很好时,大家都会很享受这段答辩时间。然而,即使缺乏积极的交流,也并不意味着应该重新准备又一次的答辩。

最后,要清楚答辩时可以携带哪些材料。答辩前肯定要向答辩委员会提交完整的论文,这时自己也要留存一份。因为经常会出现像"在某页第某行中这样说是什么意思?"这样的问题。如果允许,建议在答辩时带上统计输出的表格、量表的复印件、文章中提及的文献以及答辩中需要给出的要点提示等,可以将这些材料放在桌下或者角落里。作者在以往的答辩中曾经听到过这样的问题,"表 462 中报告的数字是否正确?""史密斯和琼斯是否真的做出了三因子而非四因子结构的量表",以及"在测量某一特定因变量时都包括哪些内容?"等。如果在回答提问时遇到上述这种需要引证的问题,事前准备的、充分的相关材料就会发挥重要作用。

模拟答辩

答辩过程十分重要,尤其是申请博士学位的答辩,因此值得多花一些时间来排

练。如果可能,最好安排一次模拟答辩,这样可以对报告和问答进行预演。同学(特别是那些已经通过答辩的研究生同学)和导师是最理想的观众,如果能说服他们拿出一点时间来参加自己的答辩预演将会十分有帮助。对一些人来说,饮少量酒也会有较好的效果!①

在模拟答辩之前,可以请导师帮忙审核一下准备报告的提纲。做些适当的修改之后再进行排练。建议进行两次模拟答辩,第一次是面对自己,用录音机或录像机录制下来,以备在模拟报告之后进行自我审查;另一次则是面对观众,在面对观众进行预演之前要用计时器记录独自彩排的报告时间,并进行调整。各学校和院系对答辩时间有不同的要求,要弄清自己有多少时间,并据此对报告进行调整。

预演之前也要做好充分的准备。准备好将要使用的视听呈现设备,并像正式答辩那样进行操作。尽可能选择与正式答辩时场景相似的场所,而且最好在与正式答辩相同的时间段内进行排练。甚至可以进行"着装彩排",穿上打算在答辩时穿的衣服等。总之,排练时各个细节与真实场景越相似越好。

在报告之后,请观众像答辩委员会成员那样提出一些问题。可以让听众提一些可能在正式答辩中会遇到的疑难问题(例如,关于理论或统计的问题)。在这个时候进行练习没有什么风险,而且可以得到一些反馈。

在模拟答辩结束之后要留出足够的时间来听取观众的反馈意见,询问自己在报告时以及回答问题时的表现。在看或者听自己的录音录像时也要注意这些问题。专栏 13.3 给出了报告及回答提问时的一些重要行为的核查表。可以考虑请观看模拟答辩的观众填写一份与此类似的清单,以此来了解答辩时听众会比较注意哪些方面。此外,还可以向听众询问关于改善答辩的建议,并进行记录。

专栏 13.3

开题与答辩时要注意的行为清单

是	否	报告部分:
□	□	1. 开篇是否适当?
□	□	2. 声音大小是否适当? 语调是否抑扬顿挫?
□	□	3. 是否有适当的微笑?
□	□	4. 与听众是否有目光接触?
□	□	5. 看起来是否充满活力? 有没有坐下或者用手支撑下颚?
□	□	6. 是否能够根据事先的准备进行演讲,而非仅仅读材料或者幻灯片?
□	□	7. 是否清晰地介绍了重点内容?
□	□	8. 报告过程中各部分之间是否有较好的过渡联接?
□	□	9. 时间分配是否适当?
□	□	10. 视听呈现是否清晰? 房间中每个角落都应该能够看清或听清。

① 少量饮酒也许会将某些人的兴奋性调节到合适的水平,但是因人而异,不可一概而论。——译者注

专栏 13.3 (续)

□ □ 11. 是否需要准备其他额外的视听材料?
□ □ 12. 是否有较多的"嗯""呃"咳嗽等中断。
□ □ 13. 事先是否有人组织,换言之,是否有人告诉观众报告将在什么时候开始。
□ □ 14. 听众是否能跟上报告的思路?
□ □ 15. 报告是否简明扼要?
□ □ 16. 是否避免了不必要的细节?
□ □ 17. 是否避免了会使自己分心的习惯性小动作(多动,抓衣服、胡子或者头发)?
□ □ 18. 是否适当地使用了专业术语?
□ □ 19. 是否清楚地解释了复杂的过程?

是 否 答问部分:
□ □ 20. 回答是否解决了提出的问题?
□ □ 21. 回答是否简洁?
□ □ 22. 问答过程中是否展示了自信?
□ □ 23. 是否进行了适当的标注(如,标明引用内容等)?
□ □ 24. 对于对立性的问题,回答是否有攻击性?
□ □ 25. 在做出反应之前,是否把握了难于理解的问题?

如果对某个特定部分比较担心(如,不确定应该如何说明研究设计或描述因素分析的结果),还可以征求听众的意见,探讨怎样能做得更好。要开放和反思,避免过度的自我保护。要记得这些听众是朋友,正在帮助你准备答辩,要听取听众的建议,并认真考虑。

建设性的意见

即使是最雄辩的发言者,在面对观众时也可能有一点儿轻微的焦虑。有些学生一想到答辩就会出现过度的焦虑。一定程度的焦虑是正常的:答辩是取得学位的最后一个重要环节,这是一种很有意义的情境。不清楚答辩时应该怎样做以及不明确答辩的规则也会导致焦虑,一位研究焦虑的专家大卫·巴洛提出的理论认为"焦虑的核心是不可预测性和不可控制性"(Barlow,2000,p. 1254)。虽然答辩既非完全不可预测也非完全无法控制,但这的确不是完全由答辩者自己掌控的过程。

正因如此,在开题或者答辩报告会上感到一定程度的焦虑是正常的。可以用倒U形的唤醒状态曲线来说明焦虑现象,中等水平的焦虑可以对表现有促进作用,有助于将注意力集中到任务上。但过高水平的焦虑会阻碍发言者独立思考,影响明确表达观点。开题或者答辩报告成功的一个关键因素是控制焦虑水平并使其发挥建设性的作用。

研究发现焦虑与3个要素相关(如,Barlow,2000;Lang,1971)。第一是认知成分,包括对焦虑的主观体验,对于情境和感受的认知。第二是生理成分,涉及焦虑的生理基础(如,心率和神经系统活动的增加、手心出汗或呼吸急促)。第三是动作,也与表现相关(如,步伐、咳嗽、目光接触等)。有效管理这3个焦虑要素有助于更

好地应对开题或者答辩报告会。

思维管理

增加“面试焦虑”的认知因素主要有两种:①对未知的恐惧(他们会问什么?我能够回答吗?);②非理性思维(我将注定会失败,在生命的其余时间里我都会失业)。如果这其中的任何一种认知方式长期持续地存在思维之中,就需要认真考虑如何应对开题和答辩。倾听自己内心深处的声音,这些声音是否围绕着对于未知或潜在的灾难性结果的恐惧?如果你内心深处的声音符合上述任意一种,请继续读下去。

如果你受到开题或答辩的模糊性以及不确定性的困扰,那么就要试着使这一过程明确起来,本章的内容应该会有所帮助。另外,可以和经历过这些过程的人讨论曾经在开题或答辩过程中发生的事情。从同学和导师处收集一些共性问题,列成清单,再添加到专栏 13.2 中去。论文答辩是对公众开放的,因此可以去本院系的答辩会进行观摩(最好选择去听最有可能成功的答辩,这样才更可能看到成功的榜样)。这些准备工作,可以降低对即将参加的开题和答辩的不确定性,会使自己觉得准备得更充分。

应该如何应对非理性思维呢?40 多年前,艾利斯(如,Ellis & Harper,1961)提出这样的观点:那些不合逻辑的想法是和抑郁、焦虑、生气及类似的不良情绪联系在一起的。与焦虑密切相关的一种非理性思维是灾难性的思考:“如果事情不能完全如愿,那将会糟透了。”另一种非理性思维则是完美主义观点:“必须要在所有场合都完美无缺地掌控自己。”这两种信念都可能会导致焦虑,而且可能同时对开题或答辩产生不良影响。

解决非理性思维的关键在于对其提出质疑,使自己的思维更符合逻辑。首先来看完美主义观点:在开题和答辩过程中一定要完美地掌控自己吗?不,事实上只需要适当地对其进行把握。反复思考对问题的回答,在理解或回答问题时遇到的偶然性失误不会导致全盘皆输。要意识到并规划自己可能存在的问题。出现一些问题很正常,开题和答辩并不要求十全十美。

那么该如何应对将产生灾难性后果的想法呢?一些学生害怕自己不能正确地回答问题。然后对自己说:“我会失败的。然后所有的工作都将会每况愈下,无家可归、失业、没有人会爱我。”这时就要更加理性地检查这些想法背后的前提假设。首先,不能回答问题的证据是什么?你知道自己做了什么,对吧?在这一点上,你应该比房间里的任何其他人更了解自己研究领域的内容,当然,可能是除了导师之外。事先进行合理的预测可以降低回答大部分问题的难度,仅剩下少数几个问题难以回答,但这也可以通过机智的临场发挥来应对。其实开题和答辩过程可能是有趣的,这一过程可能会是智力上的碰撞,与一群聪明人就自己的工作进行几小时愉快的探讨是件很有意思的事情。

一个偶然的难题会导致答辩失败么?答案通常是否定的。总的来说,简单的问题一般比较容易把握,但较难问题则不易把握。总之,过去的行为是对未来行为的最好预测,如果你在学习期间做得很好,那么一般在开题或者答辩时也会不错。即使不能回答好一或两道问题也不太可能导致完全失败。要知道,所有老师都真心希望学生成功。

如果答辩失败了呢?最坏的结果会是什么?学校一般都会允许那些通过毕业论文或学位论文的书面部分,但没有通过口头答辩部分的同学重新参加答辩。所以

最坏的情况也就是可能因为需要重新参加答辩而感到尴尬。但是尴尬无关紧要，毫无疑问终究还是会成功。尽管你可能很忙，但仍然可以利用业余时间为答辩做充分的准备，在这一点上彻底失败是不太可能的。

调节身体状态

控制焦虑的另一个关键是降低身体的唤醒水平。适当调整非理性思维可以降低唤醒水平，另外还有许多技巧性的准备工作也会有助于对生理机能做出调整，包括渐进式放松、调节、主动性训练、自我催眠、呼吸练习、生物反馈和愉快的想象等。在本章后面提供的补充资源中可以找到一些适合自己的练习方式。

对于焦虑生理因素的调节，有研究发现了一些有效的策略。比如可以通过事先将引起焦虑的情景视觉化来降低焦虑，许多运动员就是通过这种形式的想象来进行训练的。还可以在安静的场所静坐来降低焦虑，闭上眼睛，想象自己放松地坐在最喜欢的沙滩或山坡上。另外，进行全身放松也可能会有用，要有意识地放松身体的每个部分。如果觉得这些技巧不容易实现，还可以参考放松训练或机能反馈疗法的说明，许多医疗中心或者大学都提供这种说明。史密斯（Smith，2005）和戴维斯等（Davis，Robbins-Eshelman，& Mckay，2002）都著有详尽的指导性书籍。

调节身体状态的最后一个方法是调整呼吸方式，使之有助于降低而非增加焦虑。在压力大的时候呼吸会比平时快而轻，这会导致焦虑的生理反应（如，头昏眼花，手脚麻木等；Gervirtz，2002）。解决这个问题的办法是用横膈膜（肋骨下的一大块肌肉）缓慢地、深深地呼吸，可以把手放在肚子上来锻炼使用横膈膜呼吸。当吸入空气时肚子会突起，呼出空气时肚子则会凹进去。想象肚子是个气球，把它充满空气再放出去。静坐几分钟并集中精力地使用横膈膜进行规律地呼吸（尝试每分钟 5 ~ 7 次完整呼吸），这曾帮助许多人减轻了焦虑。

除了练习调节生理觉醒的技能，还要在答辩前制订合理的时间表进行练习。曾经有一个学生就是通过合理地安排练习时间，在博士学位答辩时很好地控制了焦虑情绪，表现优秀。

调节生理机能的另一种方式是使用药物。通过药物调节唤醒状态会比学习生理唤醒技巧更容易，但药物可能会有抑制认知过程的副作用。因此不建议用药理方法应对焦虑，除非能够确保不会产生消极的副作用。

管理动作行为

影响焦虑的最后一个因素是动作，这会影响你在他人面前的表现。记住，不要将动作的、生理的和认知的焦虑表现出来。因此，尽管内心焦虑得发抖，也不必在行为中表现出来。许多优秀的演讲者都承认感到内在的焦虑——但观众察觉不到这种焦虑。

想要在答辩过程中表现得既专业又熟练，需要良好的准备和练习。要记住下面的原则：事先做好缜密的计划有助于表现得更好。可以想象一些已通过答辩的同学的行为，并效仿。记住，不需要觉得自己准备很充分，只要假装准备很充分就足够了。反复练习也会相当有用。专栏 13.4 呈现了保罗（Paul，1966）的焦虑行为量表。可以以此来识别让自己看起来焦虑的表现，并试着尽量减少这些行为。还可以试着摄录自己的演讲，然后进行评价，这时你也许会为自己的冷静表现感到惊讶。事实上人们大多数时候会比自己感觉到的更好。

专栏 13.4

焦虑行为量表

等级：______

姓名：______

日期：______ 演讲编号：______ 地点：______

行为观察 时间阶段	1	2	3	4	5	6	7	8	9	10
1. 踱步										
2. 摇晃										
3. 脚步拖沓										
4. 发抖										
5. 手臂和手的动作（摇晃、抓握、摆弄等）										
6. 手臂僵硬										
7. 手的控制（在口袋里、在背后、双手紧握）										
8. 手部颤抖										
9. 缺乏眼神接触										
10. 脸部肌肉紧张（憔悴、痉挛、扭曲）										
11. 面无表情										
12. 面色苍白										
13. 脸红										
14. 嘴唇湿润										
15. 吞口水										
16. 咳嗽										
17. 呼吸粗重										
18. 流汗（脸、手、腋下）										
19. 声音颤抖										
20. 演讲停顿或口吃										

评价：______

注：表格取自 *Insight Versus Desensitization in Psychotherapy*, by G. L. Paul,

严重焦虑问题

如前所述，在口头答辩过程中感到紧张是正常甚至有益的。但如果曾在公众面前讲话时或者在社会情境中经历过严重的焦虑，则可以通过应用一些方法来控制和减少答辩中的焦虑情绪。进行发言练习也许会有助于减少焦虑（http://www.toastmasters.org）。这是一个国际性的网站，在很多城市都有其分站，你可以和最近的分站取得联系。对于一些严重的焦虑者，还可以寻求专业的帮助，在专家的指导下学习焦虑管理。一些专业性组织（如，行为与认知治疗协会，美国焦虑障碍协会等）有提供相关信息的网站。美国焦虑障碍协会还列出了一系列自助书籍。

重读论文

最后，还要强调的是要在答辩前一两周内重读论文。论文的一些章节也许是几个月甚至一两年前完成的，你是否能清楚地记得数据的评分方式？缺失值是如何处理的？被试筛选是怎样进行的？在论文中哪一部分可以找到与测量或分析相关的细节？你是否记得引发研究的关键细节，以及与本研究最相近的研究？你是否能够对这些内容进行符合逻辑的探讨？

如果对上述任何问题的答案是否定的，就需要重读所写的文字，回顾收集数据和分析数据的程序。比如，一个学生因为不能回答修订问卷的细节问题而几乎不能通过答辩。如果没有修订问卷，则可以解释研究中的异常结果，但同时也说明研究结果不可靠。但事实上，这个学生对数据进行了正确的修订和评分，但答辩时却无法回忆出修订的过程。答辩委员会考虑到其对其他问题回答得较好，才暂且允许通过，并要求其重新阅读思考修订程序，如果对问卷的修订不正确，还需要重写论文并重新答辩。然而，最关键的一点是，如果这个学生在答辩之前一周重读自己的论文并温习实验程序，就可以避免这种情况。很多学生在答辩中能轻松处理相似的问题，就是因为很熟悉自己的论文和研究。这说明认真的前期准备有助于你在答辩中表现得更好。

开题报告和毕业答辩可能令人焦虑，但也可能是愉快和收获的过程。实际上，事先进行考察和准备对研究中的每一步都是有益的，周密的准备有助于研究项目的顺利进行并最终取得成功。

补充资源

答辩过程与技巧

Pyrczak, F. (2000). *Completing your thesis or dissertation.* Los Angeles: Auttor.

焦虑管理

Anxiety Disorders Association of America. (2005). Retrieved November 17, 2005, from http://www.adaa.org/
Bourne, E. J. (2001). *Anxiety and phobia workbook*(3rd ed.). Oakland, CA: New Harbinger.
Davis, M., Robbins-Eshelman, E., & McKay, M. (2000). *The relaxation and stress reduction workbook* (5th

ed.). Oakland, CA: New Harbinger.

Greenberger, D., & Padesky, C. (1995). *Mind over mood: Change how you feel by changing the way you think.* New York: Guilford Press.

Smith, J. C. (2005), *Relaxation, meditation, and mindfulness: Self-training manual.* Morrisville, NC: Lulu Press.

Toastmaster's International. (2005). Retrieved November 17, 2005, from http://www.toastmasters.org/

思考与练习

开题和答辩管理

☐ 准备开题报告会

—了解相关程序(见专栏 13.1)

—准备记笔记

☐ 准备正式的毕业答辩

—了解相关程序(见专栏 13.1)

—准备报告

—准备视听材料

—制定回答提问的策略(见专栏 13.2)

—练习答辩过程(见专栏 13.3 和专栏 13.4)

—进行焦虑管理

—重读毕业论文终稿

14

研究发表

现在,已经敲定了毕业论文的终稿,通过了毕业论文答辩,并向图书馆提交了论文的电子版,一切看似已经基本结束。那么,此时该做些什么呢?休息么?不,至少不能将研究置之不管。从这项研究的最初构想到最终完成,你也许已经为之付出了几个月甚至几年的时间和努力,正如由基贝拉(Yogi Berra)所述,在这一过程真正完成之前不能宣告结束。从某种意义上说,研究没有终点。正如图 14.1 所示,研究是一个循环的过程,从提出问题开始,在将这一问题解决之后,又以进一步的新问题开始新的循环。事实上,每一项研究都可能为自己或者其他研究者提出新的问题。所以也可以将研究看成是解决问题的过程,而非仅仅是为了获得一个特定的结果。

那么,当完成了研究的主体部分(如,项目的某方面,或指定的论文),并同答辩委员会讨论了研究结果之后,应该在何时发表这一研究呢?最佳时机是在答辩之后马上开始行动,着手准备发表自己的研究。这个时候动机通常较强,如果错过这段时间可能就不再想整理研究了,因此,不要拖延。如果研究不能发表,或者不像图 14.1 所示的那样清晰,则可以认为这一研究过程不完整,没能真正完成一个研究周期,以至于其他人没有机会从这项研究中学习。

就像第 4 章中提到的,研究是累积的过程,每一项新的研究都建立在已有研究的基础之上。如果没有前人研究发表的文献资料,研究就会缺乏理论和实证基础,也无法向前人学习提高研究水平和论文质量的经验。此外,与他人交流分享并征求意见也有助于更客观地评价自己花费大量时间和精力进行的研究。能够通过开题和答辩的研究通常不会糟糕到不值一提,为完成研究所付出的努力通常也不只是为了获得文凭,然后将之挂在墙上。

最后,从不太崇高但更实际的方面看,报告和发表研究可能会对找工作有利。潜在的雇主可能会对精心准备的精彩报告印象深刻。如果计划毕业以后继续从事科研事业,或者想要申请研究奖学金,或者希望到研究机构中工作,那么提交过研究报告或发表过论文等可能会有助于相关申请。如果在获得硕士学位之后想继续攻读博士学位,就更要知道教师一般比较喜欢做研究积极主动并乐于报告和投稿的学生。

尽管一些硕士研究生的研究缺乏发表的机会,但还是有许多的学位论文是适合在会议中报告或者在杂志上发表的。如果你愿意与大家分享自己的研究发现,就需要探讨一下发表文章的策略:应该如何借助媒体向专业人士或者普通大众展示自己的研究?

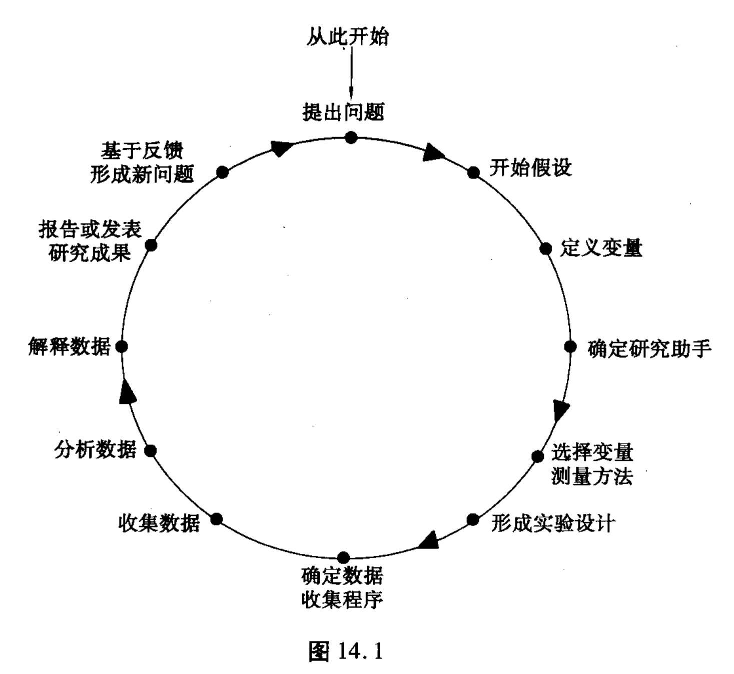


图 14.1

研究周期选自 *Evaluating Outcomes: Empirical Tools for Effective Practice* (p. 41), by J. D. Cone, 2001. Washington, DC: American Psychological Association. Copyright 2001 by the American Psychological Association.

先报告,再投稿

发表研究论文的理想进程是先向专业人员报告这一研究并且取得其反馈意见,将这些反馈意见整合到自己对研究的思考中去,然后准备一份正式的投稿稿件。同行的帮助常常十分有用。此外,通过报告自己的研究(尤其是在会议中)可以获得机会同对这项研究感兴趣的人沟通。这可以扩展自己在本专业中的人际网络,同时也有助于获得对研究的支持和鼓励。想想本专业领域中的那些重量级人物,如果他们中有人过来称赞你的研究,那岂不是很令人激动?

报告研究成果的众多方式可以看成是由非正式到正式的连续体。对研究发现最松散的非正式讨论可能是和同学边喝咖啡边进行的,最正式的则是在斯德哥尔摩诺贝尔研究所向委员会报告自己的研究。在两者之间,有院系的座谈会、找工作面试中对自己研究的介绍以及多种非正式(如,在电台或者电视节目中介绍)和正式的报告(如,在市级、地区级、国家级或者国际性专业会议上的报告)。大多数人的研究和报告都不那么正式,只有很少的心理学研究报告能够像获得美国心理学会的杰出贡献奖、总统勋章、诺贝尔奖等奖项高水平的研究报告那样正式。

许多院系会定期安排座谈会。在座谈会上会请学生、教师或者一些访问学者报告自己的研究。如果有这样的机会,则建议你在完成答辩之后尽快安排进行这种形式的报告。实际上,这是你第二次报告自己的研究了,因为在面对学位委员会答辩时,已经有过一次这样的经验。

也许有人会持这样的想法,“好吧,我会做个报告,但在本校报告没有什么意

思,因为听众不太重要,因此不会对我造成困扰。”事实上,面对熟悉的人做报告最具挑战性,因为熟悉的人比其他人更清楚这项研究的内容及存在的问题,也能够比较放松地就此进行探讨,而且不太可能对研究程序和结果等细节提出问题。也许有人会有另外的想法,“这些人每天都会遇到,如果报告很失败会怎么样呢?是否要偷偷穿过走廊或只在夜间到学校来,以避免遇到对报告表示轻蔑或怜悯的同学、同事?”这种担心是很正常的,也正因此有人认为如果可以面对熟悉的人做好报告,就可以做好任何其他报告了。

除了院系的座谈会,还可以在市级、地区级、国家级和国际性的专业会议上进行报告。决定是否向这些会议投稿时要考虑很多因素,包括会议是否接收研究报告、会议日程(如在一年中的什么时间召开会议)、费用以及会议的专业性等。如果会议接收研究报告,那么最需要考虑的因素可能就是会议的专业性问题了。如果打算毕业后直接工作,就应该考虑向有助于找到合适工作的会议提交报告。如果对工作的地域有要求,则可以考虑市级的或者地区级的会议。有时候,用人机构的与会代表会在会议过程中安排面试,并倾向于录用在会议上表现较好的报告者,为之打开方便之门。即使没有用人单位正式安排面试或提供就业机会,在市级或地区级的会议上展示自己也有助于在这一地区找到较理想的工作。

如果你倾向于继续从事学术研究,则可以考虑国家级的会议。在美国,有美国心理协会年会、APA 年会以及特定研究领域的会议(如,行为分析协会、儿童发展社会性研究协会、行为与认知治疗协会等)等,这些会议在一定程度上也具有人员安置的功能。在其主办的近期刊物和网站上可以找到会议的相关信息及时间安排。需要注意的是,要在会议召开之前很长时间就投出摘要。比如,行为和认知治疗大会在每年 11 月份召开,但通常要求在 3 月份之前就提交摘要;APA 年会通常在 8 月份举行会议,但前一年的 11 月中旬就需要投摘要。显然,在这种会议上做研究报告需要事先进行规划。这个过程非常重要,因此强烈建议将会议报告这一项目列入第 3 章中制订的研究时间表中。

当开始探索如何在会议上做报告时,会发现不同组织机构安排的报告形式有所不同。有海报介绍、口头报告、座谈会讨论、专题讨论会、讲习班、小组讨论等各种各样的形式,其中的前 3 种报告形式最为常见。较独立的研究比较适用于以张贴海报或张贴论文的形式来呈现。而如果要报告的是一个比较大的研究项目,则可以要求会议安排一个专题研讨会,围绕一个主题展开报告与探讨。如果没有报告的经验,可以参加对自己研究的专题的讨论,也可以将报告放在某一分会场中进行。

一般情况下,会议会安排经验较丰富的研究者在专题研讨会上做报告,这些研究者一般都是重大科研项目的主持人,会从自己的项目中选择一两项研究进行报告。学生的研究通常是以书面形式或海报张贴形式出现在会议上,下面将主要讨论专题研讨会和张贴海报这两种形式。

缩减研究报告

这一节强调“缩减”,是因为很多投稿论文只是将学位论文削减到可接受的字

数范围之内。字典中对缩减过程的定义十分恰当："削减掉超出的、过量的或不恰当的部分……；即削弱或减少"（韦氏在线词典）。

如果希望在会议上报告自己的研究，则必须事先提交研究摘要。这要求将上百页的内容浓缩到一两页之内，这项工作可以从论文原有的摘要出发。好的摘要应该包括研究的原理、方法、结果，并简要介绍研究的意义和影响。应该用大约60%的篇幅来介绍研究方法和成果，这部分应该说明研究中最重要的细节，因为会议委员会将根据摘要来决定一篇文章的取舍。正因如此，要明确地概述自己的研究工作，说明研究方法的合理性。专栏14.1给出了APA会议接收的一篇摘要，这篇摘要是建立在一篇学位论文的基础之上的（Ben-Horin & Foster，2001），比较典型，也许有助于读者对自己的摘要进行规划。

假设会议委员会或组织者认为这项研究很有价值，并在会议上安排了报告的时间，或者假设有用人单位安排了面试。那么也需要进行缩减工作，同样是要根据研究论文准备报告或海报，这两种缩减略有不同，但是都与第13章中介绍的准备开题和答辩报告的方法有很多共通之处。

口头报告

进行口头报告时首先要考虑应该报告多少研究内容。这取决于许多因素，比如有多少报告时间、听众可能是谁、是否可以使用视听呈现设备等一系列要考虑的问题。在准备研究报告之前要回答这些问题。一般区域性或国家性会议上的报告时间为10～20分钟，通常会提供投影仪或幻灯机。有时候也可以使用PowerPoint™或AutoPlilot来辅助报告，但需要把自己的笔记本电脑连接到显示屏上。在工作面试或在本院系的学术讨论会上，大概会给一个小时的陈述时间。在这种情况下，可以报告45分钟，以便留出一点时间解决可能出现的问题：比如报告不能按时开始、设备仪器故障等。第13章提出了许多关于口头报告的建议，这里不再重复。

在讨论答辩的部分强调过练习的重要性，在准备会议报告时练习同样重要。这只需要请人听一听你准备要报告的内容，听众可以是任何人。当然，最好是邀请导师、研究小组成员和研究生同学。也可以面对一个本科生班级进行着装彩排，还可以在其他一些组织和群体面前练习。

专栏14.1

海报介绍的摘要

受监禁妇女的精神问题和关系侵犯行为

为了理解和检测受监禁女性的精神问题，需要研究这些女性的侵犯性行为。现有研究表明，女性群体中的关系侵犯（是指损害个体与他人的关系）可能多于男性群体。此外，关系侵犯与人身攻击有许多共同之处。基于这些研究结果，推测关系侵犯可能是女性精神问题中非常重要但是被忽略了的因素。此外，评估关系侵犯可能还有助于完善受监禁女性精神问题量表。在这项研究中，还考察了监禁机构的不当行为。

专栏 14.1 (续)

本研究采用相关设计考察受监禁女性精神问题、关系侵犯、机构不当行为这三者之间的关系。这些受监禁的女性(共64人,其中包括20名非裔美国人,27名白种美国人,17名其他人种)在低、中、高的监禁程度上被关押了2个月以上。通过修订版的精神病核查量表(PCL-R)来衡量这些女性的精神问题(PCL-R的总分、因子1得分、因子2得分在不同评分者间的相关分别为0.93,0.70,0.90),评估每组中关系侵犯、人身攻击、亲社会行为的情况。通过调查这些研究对象的档案来评估机构的不当行为(评估者间一致性为1.00)。精神问题测量工具(改编自高校学生量表)有很高的信度(关系侵犯、人身攻击、亲社会行为的一致性系数分别为0.95,0.95,0.92)和效度(关系侵犯与人身攻击和亲社会行为有着显著的相关,$r=0.83$, -0.33,$p<0.05$,但与特质性焦虑量表的得分不相关$r=0.14$)。

结果表明,关系侵犯与机构的不当行为相关显著,$r=0.38$,$p<0.05$。虽然关系侵犯与精神问题评估的总分和因子分都相关,但当控制种族变量时这种相关就消失了。分别考察两种种族女性的内部相关性,尽管分组后相对较小的样本量影响做出确定的结论,但还是发现对白种人来说关系侵犯与精神问题没有相关,而非裔美国女性组则存在相关。分层回归分析表明,精神问题测量的因子2能够预测机构的不当行为,但仅限于非洲裔的美国妇女群体。与预期相反的是,关系侵犯在预测机构的不当行为与种族和精神问题的关系时,没有显著性差异。

这项研究的结果表明,关系侵犯的确是受监禁女性(尤其非洲女性)心理问题中的重要因素,但过去一直没有注意到这一因素。当前,反社会性人格障碍的评定标准和PCL-R对心理问题的评价标准都没有针对女性的项目或者内容。对关系侵犯的评估开阔了研究女性心理问题的视野,指出非暴力女性会通过关系侵犯行为进行进攻。本研究表明在研究攻击性行为、心理问题、机构的不当行为以及累犯等问题时,种族作为中介变量的重要性。

注:选自H. Ben-Horin和S. L. Foster在2001年美国心理学会年会上提交的摘要。使用经过作者授权。

海报介绍

大多数地区或国家的专业组织,比较倾向于接收个人以张贴海报的形式提交的研究报告。这种形式的好处是,可以在自己的海报旁边与对此研究感兴趣的人进行讨论。还可以在一块像布告栏一样的厚重展示板上呈现研究的核心信息。在张贴海报的会议上,需要像准备正式的研究报告或投稿稿件那样准备介绍研究的书面材料。要在海报中说明研究的各部分内容,即摘要、导言、方法、结果等,然后张贴到展示板上。海报介绍通常是围绕一个特定主题(如,人类的学习、诊断评估等)进行的。张贴者要在会议中的指定时间段(如,15时)在指定位置张贴海报。作者要站在海报附近,在与会者愿意的时候与其进行讨论。

不同会议对张贴海报的要求有所不同。一般情况下,大多数张贴守则规定:①字体大小要保证使读者在2~3英尺距离之外能够看清;②在海报附近提供研究的完整介绍;③至少一位作者要守在海报旁边,以便同感兴趣的与会者进行交流。有时候准备的介绍研究的发放材料可能会不够,这时可以考虑记下对此感兴趣的与会者的联系方式,在会议结束后再给其邮寄材料或者发电子邮件。如果打算用投递的方式邮寄,比较节省时间的方法是准备能粘贴的便签,让与会者将自己的地址写在上面①。学生会员通常可以从张贴的海报中学习如何准备海报,并可列出有助于

① 这样就可以在邮寄材料时直接将写好地址的便签贴在信封上。——译者注

筹备海报的项目清单。专栏14.2呈现了这种项目清单的模板。

决定在海报中呈现什么内容之前,要考虑到页数和字数的限制。这时会注意到,清单中的第11和12项要求海报正文最多包含250个字。对研究内容和基本原理的介绍最多不超过1页,研究方法2页,研究结果1或2页,图表或表格不超过3个,最后用1页进行讨论。清单上建议不要使用超过1个图表,这里则建议图或表格不要超过3个。有时,也可以将研究结果和讨论部分结合到一起。因为海报上的字体通常会比较大,就尤其要注意做到简洁明确。请记住,读者往往会边走边快速浏览,以了解研究的主旨,随后再挑选感兴趣的会议报告来阅读(摘要应该以最简形式的文字对研究进行概述,会议报告则要求提供研究的细节)。可以采用在摘要的基础上按重要性顺序增添细节的方法,直到达到要求的页数或者字数,以这种方式形成的海报效果很好。

专栏14.2

准备海报的核对清单

项目类别

总体__1. 研究内容清楚明确

__2. 使用 Sans Serif 字体

__3. 在图表中纳入所有研究要素

__4. 按标题顺序组织海报内容(方法、结果、讨论)

__5. 最多250字

题目__6. 标题大写

__7. 多于一行的标题也要大写

作者__8. 作者姓名完整

__9. 列出作者单位

__10. 根据单位将作者分组

正文__11. 每行最多45个字母

__12. 每段最多6行

标题__13. 所有的图表都要有标题

__14. 标题要能描述出图表的内容

图表__15. 条形图和线形图最多只能包含4组数据

__16. 线型图上的数据点清晰

__17. 图表没有留下绘图时的底线

__18. 不同的数据线使用不同的形式或颜色

__19. 条形图上的不同分类用不同的颜色表示

__20. 如果是单被试数据,可用线形图呈现

__21. 线形图要标明横纵轴标签

__22. 文字和图表要一致,互相佐证

__23. 如果使用了不同的颜色,要注意颜色对比清楚

__24. 成组的数据要报告对变异的测量(如,误差范围)

表格__25. 总共不多于1个表格

注:选自 http://www.abainternational.org/sub/membersvcs/Studentnews/Studentpres.asp。

海报的风格和形式有很多种,选择起来常常弄得人眼花缭乱。要记住,当研究者张贴海报时,不仅代表自己,也代表合作者和学校,海报的视觉形象是文章内容的重要表征。大多数人在看到手写得很差或者打印得很差的海报时会庆幸作者所在单位不是自己的学校,另一个极端是,如果海报过于鲜艳张扬,可能会掩盖研究的实质性内容。有品味、清晰、易读是评价海报的最重要标准。但要记得,研究的内容才是最重要的,所以在制作海报时也要将内容放在首位进行考虑。

成功的海报一般符合以下标准。首先,要遵循会议主办方提供的准则。这不是发挥创造性的地方,记住,这些准则都是有一定道理的。第二,简单。研究者常常要携带海报乘飞机去参加会议,而且大概只有约 15 分钟的时间来张贴或收起海报。主办方通常只提供张贴海报的展板,因此一定要记得带一些图钉或订书钉。除非有专门的部门为你制作精美的海报,否则最好制作较为简单的海报。但即使最简单的海报也要求用大号字体呈现内容,然后在页边画上 1/4 或者 1/3 英尺宽的边框,而且最好全部使用相同的背景颜色。要尽量让每一部分(摘要、简介等)的内容单独呈现在一个页面上,如果不得不分开在两页上呈现,就要考虑使用能容纳两页的展板,这有助于与会者浏览研究报告。张贴材料时通常要按照从左到右的顺序排列。如果图表、表格与其他内容的页面大小不同,则很难按从左到右的顺序进行排列,这时可以考虑使用大箭头,以指示读者按顺序阅读。

如果觉得上面的内容像是一节(仅有的一节)艺术课,还可以考虑将电子版的海报内容放在磁盘里交给打印社,请他们做出一张包含所有内容的单色大图或者大张海报。这样可以将之卷起来随身携带,在张贴海报时只需要打开纸卷,用大头针别在展板上即可。这种方法的主要缺点是成本较高,以及卷成筒状会导致邮寄的小小不便①。

报告后投稿

在向专业听众报告研究结果并取得反馈之后,就要计划将研究投稿发表。在这一点上,速度和简洁是两个重要因素。如上所述,不要把研究成果锁在抽屉里,并抛在脑后。要借助刚刚完成报告的势头,尽快将之发表。一条经验法则建议在答辩后 12 个月内报告研究成果并投稿。

选择期刊

发表研究成果,首先要选择一个合适的刊物。在这一点上,可以向导师寻求有效的建议。准备发表的研究也许是某一科研项目的一部分,这时可以将它投到一两个刊物上。研究者当然应该知道这一领域的研究通常发表在哪类刊物上,因为在整理相关文献,撰写文献综述部分时就应该注意到哪些期刊发表这一领域的研究比较多。

投稿时要注意期刊的倾向和特点,某些刊物的退稿率可能会较高。比如 APA 的一些刊物每年收到成百上千的稿件,因而退稿率大概在 80% 以上。这些刊物的编辑通常会努力找出理由来拒稿,以此来应对堆积如山的稿件。如果研究没有预期的好,或者存在明显的问题,则可以考虑投稿到不那么挑剔的或更有针对性的刊物,以避免几乎可以预见的拒绝。如果研究发现和研究结果是可靠的,而且研究者心胸比较宽广,不介意被拒稿后再转投相对宽松一些的刊物,则可以尝试一些更权威的

---

① 目前一般只需打印制作好的电子版海报即可。——译者注

刊物。本研究领域的导师或者专家一般能够评估大部分刊物接收投稿的概率。

选择好合适的期刊之后，要找一份该刊物的投稿说明。期刊的卷首或者卷末通常有这样的说明，或者可以在其网站上找到相关要求。投稿说明会介绍期刊刊登的研究种类、格式要求及投稿要求。这与行为科学的刊物要求相类似，一般的要求会与《美国心理学会写作手册》(APA,2001)的要求相一致。而且还要注意刊物所要求的研究类型，直接将稿件投到最合适的刊物将节省大量的时间。一般要60～90天才能收到刊物的反馈。可以通过仔细甄选投稿刊物来避免退稿及时间拖延。

准备稿件

像准备研究报告一样，要对研究进行修整以使之适合投稿。一些基于研究的刊物要求每篇文章最多不超过25页。主要发表文献综述的刊物(如，*Psychological Bulletin*, *Developmental Review*, *Psychological Review*, *Clinical Psychology Review*)则通常会接收长一点儿的文章。一些数据型的刊物(如，*medical publications*)则要求将字数压缩到较少的范围内。无论哪种情况，投稿时都要对毕业论文进行缩减。缩减到什么程度，则要看所要投稿的期刊中文章的平均长度。假如是根据APA要求撰写的文章，一般可以通过将文章的总页数(包括标题页、参考文献页等)除以3，来得到将之刊登在期刊上的页数(APA,2001)。

还有很重要的一点需要注意：不要仅仅对毕业论文进行剪切粘贴，这会导致文章连贯性差。如前所述，一个方法是根据海报内容或文章摘要来组织内容，而不要根据原有论文删减。写投稿文章只需先将研究的主要内容和细节纳到其中，然后添加次重要的内容，然后再次……直到基本达到预期的长度。

在写作的同时，要考虑投稿期刊的特别要求及典型的文章内容。对理论部分、方法部分、讨论部分有多重视？对文章进行调整，尽量使之符合期刊的这些隐性或者显性的特殊要求。

还要注意每一个细节，如果自己都不能理解自己文章中的一个段落或句子，那么对从没听说过这一研究的人来说就更不可能清楚地明白了。可以找人来读准备投稿的文章，指出不恰当的地方或不必要的细节，以使文章表达更加清楚。另外，要严格遵循APA写作要求。不严谨的文章可能会引发对研究严谨性的怀疑。

可以与导师或其他老师讨论如何完善写作。通常，学生先写一个初稿，然后请导师修改并做出点评，再据此做出调整。制订写作计划无疑会有助于及时完成初稿。

确定作者

“心理学家与著作成果道德准则”的8.12标准规范了著作权。通常投稿的论文是由学生及其导师联合协作完成的，通常以学生为第一作者。像APA写作手册(APA,2001)指出的那样，作者应该是指对论文研究做出了重大科学贡献的人，而非指进行文字写作的人。一般而言，“重大专业贡献”的人所指的是那些提出研究问题及假设，完成论文主体部分，进行统计分析，做出研究设计，提供个人收集到的数据，或安排被试等工作的研究者。对于相对较小的智力贡献和其他形式的非智力的贡献(如，收集数据，对数据统计提出意见等)，可以在脚注里表达感谢。一般以个人贡献的大小来确定作者的序位，贡献最大者为第一作者。也存在极少数的例外，“在基于一个学生的博士论文进行的一系列合作性研究中，这个学生都应被列为主要作者”(APA,2002,p.1070)。

同时,APA 伦理委员会表示,如果是由导师指定了主要变量,对研究做出重要解释,或提供了数据库,则一定要将导师列为第二作者。委员会还指出,如果是由导师指出研究领域的范围,并较多地参与研究设计和相关措施的制定,或者在投稿时做了实质性的文字工作,则可以将导师放在第一位。但如果导师只是在学生行文时提了一些小建议,鼓励学生,提供设备及经济支持,或仅做出评价等,则不可放在首位(APA 伦理学会,1983)。对于这种支持帮助,可以在脚注中声明。费恩和库尔德可(Fine & Kurdek,1993)更详细地讨论了如何确定作者及其序位。肖卡克、法提斯和伯特斯丁(Shawchuck,Fatis & Breitenstein,1986)也提出了安排作者位序的一系列步骤。

要按作者对本研究的贡献而非资历深浅来对作者进行排序(例如所有儿童都是由第一位作者观测的)。很显然,根据对研究贡献的重要性给作者排序是有必要的。因为经常有资历相对较浅的作者对研究做出更大的贡献,尤其是在他们的博士论文研究中。

有时候很难确定各个作者在共同努力完成一项特定研究时所做出的贡献大小。因此,针对作者以及作者序位的争论是很常见的,而且他们一般都不会太满意。一个好的解决方法是:在研究的初始阶段就要决定好谁该做什么,以最后做出的贡献来决定每个人最后在作者中的位置。

事实上,投稿是研究过程中的一个重要环节,将所研究的问题抛在一边不发表是不明智的。尤其是毕业论文,学生和导师都在毕业论文的研究上花费了很大的心血,而且需要进行几次书面和口头的报告,因此最好发表这一研究。在确定研究问题后,就要考虑与导师建立正规的协议(也许是书面的),包括投稿计划。这一协议应该包括论文形式,完成论文之后多少天内投稿(一般是 12 个月内),学生、导师和参与研究的其他人的责任,以及作者序位等。这份协议还要考虑到其他人中途参加研究可能造成的影响。事先制订清楚的书面协议有利于后期处理各方关系,也有利于保证研究向着发表的方向进行。

投稿

当准备把稿件投寄到事先选好的刊物的时候,要再次查阅这一杂志的投稿指导手册,以了解并遵守投稿程序。如果要将稿件发给期刊的编辑,就要找出当前编辑的姓名和地址,一些刊物的编辑大概 3 ~ 5 年更换一次。根据期刊当前的要求投寄相应份数的文章。注意稿件要完整、可读、页序排列正确。不要用订书钉将稿件钉在一起,可以用小夹子。要在稿件前面的封皮页上声明此文章并未同时投向其他刊物。如果该刊物提供匿名的点评,并且你希望能够获得这些评语,则要在信中注明。信中还要提到研究中对待被试(无论是人还是动物)符合 APA 相关研究的伦理标准。如果稿件中涉及有版权的材料,要先取得使用许可。最后提供自己的电话号码和联系地址,如果在稿件审查过程中地址有所变动,要及时通知处理这一稿件的编辑。

大多数刊物会把文章分派给在相关研究领域较有经验、手头工作相对较少的助理编辑。然后助理编辑找 2 ~ 5 个审稿人来对这篇文章做出评价,一般这些人中会有一个以上是该研究领域中的专家。这些审稿人会给出对文章的优点和不足的评议,并提供该刊物是否应发表此文章的意见。大多数编辑会书面告知收到了投稿,有时还会告知具体是哪位编辑在处理这一稿件(如,编辑或某个编辑助理,编辑助理有时被称为“执行编辑”),以及可能收到回复的时间。一般这一过程要 90 天左

右,但如果该刊物接收到的稿件较多则可能有所延迟。

如果在告知的回复时间之后一两周还没收到答复,可以打电话或者写信给负责审阅文章的编辑,礼貌地询问情况。助理编辑在找审稿人时可能会遇到困难,审稿人可能有所拖延,或者助理编辑有所拖延。以两周的间隔礼貌地重复询问一般可以促使较慢的编辑更及时地完成对文章的评审过程。

即使收到编辑的反馈,也还没有结束研究工作。期刊的编辑一般会给出4种答复:最罕见的是无需修改,直接发表;而最常见的是拒稿,这时会提供一些匿名审稿人的意见;第三种是在按要求进行一些修改后予以发表;第四种答复则根本不做出决策,只说编辑和审稿人难以抉择,文章有很多优点,可同时也存在很多在本文基础上通过修改也难以解决的问题。编辑会鼓励作者根据审稿人的意见重新思考行文,然后再投稿。在这种情况下,编辑会像处理新接收的稿件那样来处理新的文章,重新开始整个审稿流程。

要求修改的评审意见常常会打击初次投稿者,他们常常认为文章最后会难逃一劫。事实上,如果是拒稿率较高的较好期刊给出这种答复,可以看成是好消息,这意味着稿件可以得到又一次的机会。如果觉得可以在合适的时间内根据意见对文章进行修改,则最好再投给原来选定的刊物。如果评审意见与自己的观点有较大冲突或者需要较长时间修改,可以考虑修改后投稿到其他刊物。注意,如果根据审稿人和编辑的意见对文章进行调整,则每次投稿被接收的可能性都会增加。

鼓起勇气,甩掉负担,重新开始

因为投到最好刊物的稿件中大概有80%会被拒,所以第一次投稿时要做好被拒的准备。即使没有被拒,也很可能会受到较多批评。初次投稿者常常会对此有愤怒或者抑郁等悲观消极的反应。在这一点上要好好把握,保持平常的心态和积极的态度。将拒稿看成再次回到游戏中的一个机会,并且可以以更高的姿态来参与游戏。在理解审稿人意见时,要注意以下几点:第一,编辑要求审稿人指出稿件应该如何改进,这可以被看成是一种批评,但通常是会指出如何完善文章的建设性意见;第二,对优点的评价通常会比较简短,而对不足的说明则较多;第三,评审一般强调批评而非赞扬。由于这些原因,接到的评审意见一般都集中在文章的弱点。即使是发表在最好的刊物上的文章通常也要根据审稿人意见进行两次修改。点评是针对文章的,不要认为评论是对人的。

也正是因为初次投稿被拒率较高,所以要在答辩之后尽快撰写投稿文章。拒稿带来的坏消息是也许需要一试再试,一次又一次地修改后再投出。好消息则在于拒稿时通常会提供本领域专家的精彩评论,这有助于完善文章和研究。如果认真对待这些评论并做出合适的改进,会大大增加下一次投稿被接收的可能性。更重要的是,这些专家的意见可以提高自己批判性思维的能力及将来在这个或其他领域进行研究的能力。

顺便说一句,不要不加修改就将稿件投到另一个刊物。不同的刊物有时会请相同的审稿人,这些辛勤工作的研究者在看到一份没有根据修改意见进行修改的稿件再一次被投出时,一般不会采取善意的态度。同行审查过程是科学发展的一个重要机制,是值得所有人尊重的。

如果第一次投稿被拒(这是很可能发生的),研究者可能会倾向于把稿件丢在一边,置之不理,这种反应很自然。但是一些审稿人的评审意见只是基于对研究工作或研究发现的误解,还有可能只是对稿件做出批评性的攻击。因此很可能会得出

这样的结论,“这项研究是有缺陷的,不值得出版或发表”。要避免这种想法。要记得,学位委员会认可这项研究并承认研究的贡献。也就是说,这项研究一定是有一些价值的。大多数精心构思和谨慎进行的研究工作都值得与专业人士分享。当收到评审意见时,要认真阅读,还可以坐下来和导师进行探讨。在大概一星期左右以后再花时间来进行讨论。期间要冷静下来,将评审意见放在一边,直到再次与导师会晤前再拿出来看一看。这样做可以提供新的、较少受抵触情绪影响的观点,将能够根据评审意见思考出一些修改文章的思路。当与导师讨论的时候,可以确定如何根据修改意见逐条修改,并考虑各条之间的关系。如果不想再投稿到第一次投稿的刊物,还要想好把修改过的文稿投到哪里,并了解这一刊物对投稿的要求。

考虑将一个大项目划分为多个小项目

现在即将要把研究投稿到专业期刊。对一篇毕业论文来说,很重要的一点是能提供多种学术产物。事实上,你可能从一开始就制定了一个投稿计划。也许是计划将整体研究中的文献综述部分进行润色和修改,投到讨论类或者综述类刊物,比如《心理公告》(*Psychological Bulletin*);另外,也可能在研究中发展出一种研究设备或者计算机软件等研究工具,那么就可以单独用一篇文章进行详细介绍,然后再投到适合的刊物上去;还可能在研究中总结出某些干预方式的操作手册,具备相关的细节,那么可以出版一本手册类的小书;或者可能得到关于某种新研究设备的使用信息,也可以单独形成一篇文章进行介绍,等等。所有这些投稿都源自于一项研究的主体,但是可以单独成文投稿。

如果考虑要从一项研究中提炼出多篇论文来发表,就首先要了解 APA 关于重复出版的警告(APA,1983,2002)。重复出版,包括那些“在不同刊物或者著作中,存在全部或部分内容上重复的现象”(APA,1983,p. 167)。例如,如果要将一篇论文的文献综述部分提出来单独投稿,则必须不同于整体研究投稿时的前言或者其他部分。在投稿时要向编辑提供从同一研究拆分出来的相关文章的信息,以便编辑能有机会了解是否存在重复出版这一问题。

请注意,这里谈论的是投稿论文要与整体研究的部分内容相独立。不主张“将从同一组数据中得出的研究结果拆分成几篇研究报告来投稿”(APA,1983,p. 168)。这种做法往往会构成重复出版(因为存在大量重叠)。但如果是较长期的研究(如,对个体发展的纵向研究),则可以根据同一组数据提炼出多篇投稿文章。建议阅读《APA 写作手册》(2001)以了解更多这方面的内容。

权衡在大众媒体刊出的利弊

到现在为止,我们的讨论一直限定在向专业人群报告或呈现研究结果,但一般的公众也可以成为合适的读者。研究者将自己的研究呈现给整个学术群体,不仅有益于自己研究领域的发展,还有益于自我发展。假设你所进行的是关于行为金融学这一较新领域的研究(行为金融学研究情绪和认知对投资者决策行为的影响)(如,Shiller,1999),而且已经完成了毕业论文,发现在假设继承了大笔钱时,不同个体的投资行为存在很大的差异,有的会进行很多种投资,而另外一些人只投资于一两个项目,而且发现投资倾向与个人的风险评估存在正相关。那么就可以通过大众媒体告知公众这一研究结论,并谨慎地指出这一研究对个人投资行为的意义。某些主流

报纸的经济板块也可能会引用这一研究的结论，将研究者称为这一领域中的专家，这会有助于你开启个人投资行为研究的心理学事业！

如果你希望投身学术事业，也可能会受益于在大众主流媒体刊登自己的研究。假设研究需要地方机构、学校等的合作，那么如果报纸等媒体报道过自己的研究，如"布莱克博士的研究令人着迷，他最近被认为是 O. U. K. D 的专家助手……"这种报道将有助于开启合作之门。这里所谈到的当地报纸和电台以及电视访谈节目等进行的报道，通常不会涉及研究的全部主要结果。还有其他一些很受欢迎的研究发表渠道(如，杂志或书籍)，但是，如果研究的主体部分已经发表过，则可能不适合再次发表。

要注意，这些媒体的审稿人一般不会是本专业的专家，因此，通常不会太重视研究的学术价值。而且，要在大众媒体中报道，会很难在内容上进行把握。谨慎的措辞、严密的推理解释可能被看成是过于小心的表现，大众媒体为了追求新闻价值，常常倾向于夸大研究。一定要要求在出版前先亲自过目大众媒体准备好要报道的稿件，尤其是当文中直接引用某些语言或者观点的时候。最后，还要考虑一点，大众媒体上的报道也可能会引发重复出版的问题，可能使得专业刊物不再发表相关的研究，因此要仔细权衡。可以制订一项出版计划，纳入多种刊物，还可以纳入专业和非专业的多种媒介渠道。如果决定通过媒体向大众介绍自己的研究，则在向公众报道时也一定要遵守 APA 的结构要求。如果要包括研究的顺序和时间安排，可以考虑使用第 3 章中介绍的项目规划软件。

发表研究是研究周期的最后一步也是最重要的一步。不要因为逃避劳动和责任的自然倾向就将研究扔到旁边，去匆匆开始新的生活。每个研究者的研究都是受益于其他研究者的报告或者发表的文章，并从其中学习，你现在也要为其他人提供学习的机会。而且，能把自己的研究论文发表在刊物上也是值得庆祝的事情。所以，努力去发表研究吧，看到自己的名字变成铅字出现在刊物上是很令人激动的！

补充资源

一般出版指南/建议

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

McInerney, D. M. (2001). *Publishing your psychology research: A guide to writing for journals in psychology and related fields.* Thousand Oaks, CA: Sage.

研究伦理

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, 57, 1060-1073.

准备海报张贴的建议

Association of Behavior Analysis. (2005). *Student committee: Poster presentations.* Retrieved November 18, 2005, from http://www.abainternational.org/sub/membersvcs/Studentnews/Studentpres.asp

University of Washington School of Public Health. (2005). *Creating a poster using MS PowerPoint.* Retrieved

November 18,2005,from http://depts. washington. edu/mphpract/ppposter. html

思考与练习

研究发表

□提交会议报告
—同导师商讨并确定作者
—明确会议主题及投稿时间
—提交摘要

□准备会议报告
—明确口头报告的形式及长度
—明确张贴海报时的页数限制
—制作清晰、可读性强的海报
—准备与海报相应的宣传单
—准备好张贴海报时可能会用到的大头针、胶水等工具
—准备记录地址的便贴条,以便以后联系

□研究投稿
—选择一个刊物
—仔细阅读投稿指南
—写稿
—投稿
—准备好接受批评,但是不要被打倒

□将一个大的研究分成多篇稿件

□考虑大众传播媒体

附录:心理学研究中的伦理标准选摘

2.05 对他人的工作授权

心理学家向雇员、督导、教学研究助手或者翻译等雇佣服务人员授权工作时应注意的问题:

①避免将工作授权给存在多重关系的人,避免因此而引起谋私行为或有失客观的情况;

②只授权给受过教育、训练或具有经验、能独立或在合理的监督下胜任工作的人;

③进行监督以使其称职地工作。(也见执行标准2.02,提供紧急服务;3.05,多重关系;4.01,保密;9.01,评价基础;9.02,使用评价;9.03,评估的准许;9.07,不合格的人进行的评估。)

3.04 避免伤害

心理学家应通过合理的步骤避免使其咨询客户、病人、学生、督导、研究被试、组织客户以及其他相关工作人员受到伤害,并且尽量将不可避免的和可预见的伤害最小化。

3.08 探究关系

心理学家不应对其监督的或被其授权的人,如咨询客户、病人、学生、辅导人员、研究被试和雇员进行过分的探究。(也见执行标准3.05,多重关系;6.04,费用和资金安排;6.05,与客户/病人交易;7.07,与学生和指导者的性关系;10.05,与当前的咨询客户、病人的亲密性关系;10.06,与咨询客户/病人的其他相关或重要他人的亲密性关系;10.08,与前咨询客户/病人的亲密性关系。)

4.01 保密

心理学家最基本的义务是对其从任何媒介获得或保存的信息采用可靠的保密手段,了解由法律或相关机构对专业领域或具体关系规定的保密范围。(也见执行标准2.05,对他人的工作授权。)

4.02 讨论保密的局限性

A. 心理学家与具有技术或专业关系的个人(包括在可行的范围内,不能按照法

律法规隐瞒信息的人及其律师)和组织就:①相关的保密局限;②在心理活动中产生的可预见信息(也见执行标准3.10,知情同意)进行讨论。

B. 只有在不能处置或处置不当时,才可以讨论发生在上述关系之外的以及随后在新情况下出现的相关保密内容。

C. 通过电子手段提供服务、产品或者传输信息给客户或病人的心理学工作者,需要知会其客户或病人这些服务手段在隐私和保密性上存在局限和风险。

4.04 最小化对隐私的干扰

A. 进行写作、口头报告和咨询的心理学家应只关注交流中与研究目的有关的信息。

B. 心理学家可以出于技术或者专业的目的与了解以上保密内容的人讨论其在工作中获得的保密信息。

4.07 为了教诲和其他目的使用保密信息

心理学家不能在其作品、讲座或者其他公开媒体上泄露涉及咨询客户、病人、学生、研究被试、组织客户,或者在其教学和工作中获得的需要保密的个人身份识别信息,除非:①合理地掩饰了个人或组织的身份;②个人或组织有书面同意书;③个人或组织授予其正式的权利。

5.01 避免伪造或者欺骗惑众

A. 公开声明包括但是不只局限于付费或是免费广告、产品许可、补助金申请、驾照申请、其他资格申请、插图册、印刷品、目录表、个人简历,或者用在出版、电子传输媒体上的评论,以及法律文件、演讲和公开的口头报告中的材料。心理学家使用虚假的公开材料视为欺骗,是对自己的研究、实习、工作以及个人、组织或其隶属机构作假。

B. 心理学家不能在其受到的:①训练、经验或能力;②学历;③资格;④隶属的机构;⑤进行的服务;⑥技术或者临床基础,或研究结果及成功的程度;⑦研究费用;⑧作品或是研究发现上进行欺骗,造假。

C. 心理学家只能在其资格是:①从认证的教育机构获得的;②或是在其实习的地区得到的心理学执照时,才可以在提供的服务时声明自己具有资格。

6.01 专业和技术工作的文件以及记录的保护

心理学家建立和交流的记录要在其的控制之下,进行保护、传播、储存、保留以及管理。与专业和技术工作相关的数据应该用来:①为之后的或者其他专业的研究者提供便利条件;②允许复制其研究设计和数据分析;③符合相关学院的要求;④确保账单和付款的准确性;⑤确保得到法律的允许。(也见执行标准4.01,保密。)

6.02 保护、传播和处理专业和技术工作的保密记录

A. 无论在书面的或是使用其他媒介的情况下,对保密记录的建立、储存、使用、转移和处置要在心理学家的控制之下进行。(也见执行标准4.01,保密;6.01,专业

和技术工作的文件以及记录的保护。)

B. 如果授权给心理学家的保密信息在数据库或是记录系统中,而这一数据库或者系统对无需授权即可使用这个系统的人开放的话,心理学家需要设置密码或者通过其他技术来保护这些个人信息。

C. 当对数据进行转移和处理时,心理学家需要提前制订恰当的计划来保护记录和数据。(也见执行标准3.12,心理学服务的终止;10.09,咨询的终止。)

8. 研究和出版

8.01 机构许可

当需要进行机构审核时,心理学家应该就自己的研究目的提供精确的信息,以获得可以进行研究的早期许可。之后的研究必须与获得许可的申请草案相一致。

8.02 研究的正式许可

A. 当得到执行标准3.10中规定的正式许可时,心理学家应该告知被试以下信息:①研究目的、研究的预期长度和过程;②被试从研究中退出的权利;③拒绝或退出的可预见后果;④可能影响被试意愿的因素,如暂时的风险、不适或是有害的效果等;⑤任何潜在的研究受益;⑥保密的局限;⑦被试的奖励;⑧被试权利和研究对象。要对被试的提问做出回答。(也见执行标准8.03,对研究进行录音和采像的许可;8.05,研究正式许可的颁发;8.07,欺骗性研究。)

B. 在进行实证处理的干涉研究时,被试应该在研究进行前就清楚:①实证处理的特点;②控制组可以使用和不能使用的服务;③处理和控制组的任务手段;④如果被试临时拒绝参与或在研究开始后退出的替换处理;⑤被试的补偿和费用,是由主试支付还是由第三方支付。(也见执行标准8.02a,研究的正式许可。)

8.03 对研究进行录音和采像的许可

心理学家在收集被试的声音、图像等数据前要获得许可,除非:①研究是在公共场所进行的完全的自然观察,研究数据不会暴露个人信息或是对被试造成伤害;②研究设计包含欺骗。但是研究记录只能在得到许可的情况下使用。(也见执行标准8.07,研究中的欺骗。)

8.04 咨询客户、病人、学生和被试

A. 当心理学家对咨询客户、病人、学生或被试进行研究时,应该采用合适的方法来避免潜在的被试退出研究时造成的不利后果。

B. 当研究被试是为了课程设置或获得学分而参与研究时,被试要有选择研究项目的权利。

8.05 正式的研究许可的发放

只有以下两种情况才能获得正式的研究许可:

A. 研究不会对被试造成压力或伤害,包括,①教学设计中的普通的教育练习、课程或是教室管理方法;②使用匿名问卷、自然观察或是不会引起犯罪、公民责任问题以及损害他人的财产、就业、声誉和隐私的档案研究;③对工作和机构设置中的组织效果的研究不能影响到被试的就业和隐私。

B. 已获得法律或者联邦和有关机构的许可。

8.06 诱导研究被试

A. 心理学家应该尽量避免对研究被试提供极端的诱导或金钱上的诱惑,特别是在这些诱惑可能对被试造成压力的时候。

B. 当提供给研究被试的诱惑是心理学的专业服务时,心理学家要清楚所提供服务的特点和风险,责任和局限。(也见执行标准6.05,与咨询客户、病人交易。)

8.07 研究欺骗

A. 除非心理学家认为使用的欺骗技术具有潜在的技术、教育或是应用价值方面的显著效果,而且没有可替换的非欺骗程序,否则不能进行包含欺骗的研究。

B. 心理学家不能对被试隐瞒研究可能引起的生理疼痛或是情绪障碍。

C. 心理学家应该在合适的时候尽早向被试解释欺骗对于整体设计的作用,以及因为被试的参与而得出的更适当的结论。时间不能晚于得出研究结论之后,而且被试有权收回自己的研究数据。(也见执行标准8.08,汇报。)

8.08 汇报

A. 心理学家要告知被试研究的特点、结果和结论,而且应该采用恰当的方法消除被试可能出现的误解。

B. 如果研究技术有违人道价值观,心理学家应该使用合理的方法来减小伤害的风险。

C. 当心理学家意识到研究进程已经伤害到被试时,需要立即采取合理的手段,尽量减小这种伤害。

8.09 人文关怀和动物实验

A. 心理学家获得、照顾、使用和处置动物要符合现有的管理与规范,并且具有统一的专业标准。

B. 心理学家使用研究方法训练和照顾实验动物时,应该全程监督动物的活动,要保证考虑到动物的舒适性、健康和是否得到了人道待遇。

C. 心理学家要保证自己和所有助手都了解关于研究方法、动物的照顾、饲养和品种的说明,以及对动物的适当使用程度。(也见执行标准2.05,对他人的工作授权。)

D. 心理学家要有效减少动物的不适感、感染、疾病和疼痛。

E. 只有在没有替换程序或者研究目的对技术、教育或应用有极大价值的情况下,才可以进行让实验动物产生疼痛、压力或是丧失器官的研究程序。

F. 心理学家要在遵守技术规范并进行麻醉的情况下手术,要在术中和术后尽量

减少动物的疼痛。

G. 当需要终结动物的生命时，心理学家应尽快根据程序进行处理，并且尽量减少动物的疼痛。

8.10 报告研究结果

A. 心理学家不能伪造数据。（也见执行标准 5.01a，避免伪造或者欺骗惑众。）

B. 如果心理学家在其发表的数据中发现巨大的错误，要合理地通过修正、撤回、勘误或是其他可行的公开方式纠正错误。

8.11 剽窃

心理学家不可以把他人的全部或部分工作以及数据窃为己有，即使是引用率很低的工作或者数据也不可以。

8.12 出版

A. 心理学家只对其实际参与的或是做出贡献的工作享有责任和权利，包括著作权。

B. 无论职称如何，主要著作和其他作品反映的是其个人在相关技术或专业上的贡献。纯粹的官衔（如院长），不能代表其就享有作品的著作权。对研究或发表的作品具有次要贡献的人应该被适当的提及，如放在脚注或是介绍的部分。

C. 除了例外情况，在任何多作者文章中被列为主要作者的学生都应该是因为这篇文章是基于其博士论文写作的。指导教师和学生应该在研究进行和文章发表的过程中尽早讨论文章的著作权问题。（也见执行标准 8.12b，出版。）

8.13 重复使用已发表的数据

心理学家不能将之前已经发表过的数据，作为自己独有的研究数据来重复使用。但是可以在对数据进行正确合理的补充之后再使用这些数据。

8.14 分享已证实的研究数据

A. 在研究结果发表之后，心理学家不得拒绝把得出结论的数据提供给其他专业人士进行重新分析，使其可以对研究进行核查。除非法律要求，否则需要保护被试的隐私。这不妨碍心理学家要求有此需要的个人或者组委分担信息保护费。

B. 只能根据其声明的目的使用从其他心理学家那里得到的数据，即重新分析核查对方的研究。提供数据的心理学家对数据在其他方面的使用享有优先权。

8.15 评阅

为报告、出版、获得批准或研究的目的而进行评阅的心理学家要尊重提交者的隐私和所有权。

参考文献

Aiken, L. S. , & West, S. G. (1996). *Multiple regression: Testing and interpreting interactions.* Thousand Oaks, CA: Sage.

Ambady, N. , & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, 111, 256-274.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychological Association Ethics Committee. (1983). *Authorship guidelines for dissertation supervision.* Washington, DC: Author.

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, 57, 1060-1073.

Anastasi, A. , & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Andrews, F. M. , Klem, L. , Davidson, T. N. , O' Malley, P. M. , & Rodgers, W. L. (1981). *A guide for selecting statistical techniques for analyzing social science data* (2nd ed.). Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan.

Antony, M. M. , Orsillo, S. M. , & Roemer, L. (2001). *Practitioner's guide to empirically based measures of anxiety.* New York: Kluwer.

Anxiety Disorders Association of America. (2005). Retrieved November 17, 2005, from http://www. adaa. org/

AskOxford. Com. (n. d.). Retrieved March 7, 2005, from http://www. askoxford. com/concise _ oed/dissertation?view = get

Association for Behavior Analysis. (2005). *Student committee: Poster presentations.* Retrieved November 18, 2005, from http://www. abainternational. org/sub/membersvcs/Studentnews/Studentpres. asp

Association for Support of Graduate Students (1995). Tips for writing a literature review. *Thesis News*, 4, 8-9.

Bailer, J. C. , & Mosteller, F. (1988). Guidelines for statistical reporting in articles for medical journals: Amplifications and explanations. *Annals of Internal Medicine*, 108, 266-273.

Bakeman, R. , & Gottman, J. M. (1986). *Observing interaction: An introduction to sequential analysis.* Cambridge, England: Cambridge Univetsity Press.

Barlow, D. H. (2000). Unraveling the mysteries of anxiety and its disorders from the perspective of emotion theory. *American Psychologist*, 55, 1245-1263.

Barlow, D. (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed.). New York: Guilford Press.

Barlow, D. H. , Allen, L. B. , & Choate, M. L. (2004). Towards a unified treatment for emotional disorders. *Behavior Therapy*, 35, 205-230.

Barlow, D. H. , & Hersen, M. (1984). *Single-case experimental designs: Strategies for studying behavior change* (2nd ed.). Elmsford, NY: Pergamon Press.

Baron, R. M. , & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.

Beck, A. T. , Rush, A. J. , Shaw, B. F. , & Emery, G. (1979). *Cognitive therapy of depression.* New York: Guilford Press.

Bellack, A. S. , & Hersen, M. (1977). Self-report inventories in behavioral assessment. In J. D. Cone & R. P. Hawkins (Eds.), *Behavioral assessment: New directions in clinical psychology* (pp. 52-76). New York: Brunner/Mazel.

Bell-Dolan, D. J. , Foster, S. L. , & Sikora, D. M. (1989). Effects of sociometric testing on children's behavior and loneliness in school. *Developmental Psychology*, 25, 306-311.

Bem, D. J. (1995). Writing a review article for *Psychological Bulletin. Psychological Bulletin*, 118, 172-177.

Ben-Horin, H. , & Foster, S. L. (August, 2001). *Psychopathy and relational aggression among incarcerated women.* Paper presented at the annual meeting of the Amerian Psychological Association, San Francisco.

Billingsley, F. , White, O. R. , & Munson, R. (1980). Procedural reliability: A rationale and an example. *Behavioral*

Assessment, 2, 229-241.

Bloom, M., Fischer, J., & Orme, J. E. (2005). *Evaluating practice: Guidelines for the accountable professional* (5th ed.). Boston: Allyn & Bacon.

Bolker, J. (1998). *Writing your dissertation in 15 minutes a day*. New York: Henry Holt.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, 110, 305-314.

Bourne, E. J. (2001). *Anxiety and phobia workbook* (3rd ed.). Oakland, CA: New Harbinger.

Bruning, J. L., & Kintz, B. L. (1987). *Computational handbook of statistics* (3rd ed.). Glenview, IL: Scott, Foresman.

Burka, J., & Yuen, L. (1983). *Procrastination: Why you do it, what to do about it*. Cambridge, MA: Da Capo Press.

Butcher, J. N., Graham, J. R., Haynes, S. N., & Nelson, L. D. (Eds.). (1995). Special issue: Methodological issues in psychological assessment research. *Psychological Assessment*, 7, 227-413.

Cairns, R. B., & Green, J. A. (1979). How to assess personality and social patterns: Observations or ratings. In R. B. Cairns (Ed.), *The analysis of social interactions* (pp. 209-226). Hillsdale, NJ: Erlbaum.

Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, and discriminant validity. *American Psychologist*. 15, 546-553.

Campbell, D. T., & Fiske, D. (1959). Convergent and discriminant validation by the multitraitmultimethod matrix. *Psychological Bulletin*, 56, 81-105.

Carr, J. E., & Burkholder, E. O. (1998). Creating single-subject design graphs with Microsoft Excel. *Journal of Applied Behavior Analysis*, 31, 245-251.

Christensen, T. C., Barrett, L. P., Bliss-Moreau, E., Lebo, K., & Kaschub, C. (2003). A practical guide to experience-sampling procedures. *Journal of Happiness Studies*, 4, 53-78.

Clarke, G. N., Lewinsohn, P. M., Hops, H., & Seeley, J. R. (1992). A self-and parentreport measure of adolescent depression: The Child Behavior Checklist Depression Scale (CBCL-D). *Behavioral Assessment*, 14, 443-463.

Cleveland, W. S. (1995). *Visualizing data*. Summit, NJ: Hobart Press.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, 7, 249-253.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

Cohen, P., Cohen, J., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cone, J. D. (1977). The relevance of reliability and validity for behavioral assessment. *Behavior Therapy*, 8, 411-426.

Cone, J. D. (1978). The Behavioral Assessment Grid (BAG): A conceptual framework and a taxonomy. *Behavior Therapy*, 9, 882-888.

Cone, J. D. (1992). Accuracy and curriculum-based measurement. *School Psychology Quarterly*, 7, 22-26.

Cone, J. D. (1999a). Observational assessment: Measure development and research issues. In P. C. Kendall, J. N. Butcher, & G. N. Holmbeck (Eds.), *Handbook of research methods in clinical psychology* (2nd ed., pp. 183-223). New York: Wiley.

Cone, J. D. (Ed.). (1999b). Special section: Clinical assessment applications of Selfmonitoring. *Psychological Assessment*, 11, 411-498.

Cone, J. D. (2001). *Evaluating outcomes: Empirical tools for effective practice*. Washington, DC: American Psychological Association.

Cone, J. D., & Dalenberg, C. (2004). Ethics concepts in outcomes assessment. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (3rd ed.). Mahwah, NJ: Erlbaum.

Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: Sage.

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: HarperCollins.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements*. New York: Wiley.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathlolgy. *Journal of Consulting Psychology*, 24, 349-354.

Darley, J. M., Roediger, H. L., & Zanna, M. P. (Eds.). (2003). *The compleat academic: A practical guide for the beginning social scientist*. Washington, DC: American Psychological Association.

Davis, G. B., & Parker, C. A. (1979). *Writing the doctoral dissertation: A systematic approach*. Woodbury, NY: Barron's.

Davis, M., Robbins-Eshelman, E., & McKay, M. (2000). *The relaxation and stress reduction. workbook* (5th ed.). Oakland, CA: New Harbinger.

Delwiche, L. D., & Slaughter, S. J. (2003). *The little SAS book* (3rd ed.). Cary, NC: SAS.

DeMaster, B., Reid, J., & Twentyman, C. (1977). The effects of different amounts of feedback on observer's reliability. *Behavior Therapy*, 8, 317-329.

Dent, C., Galaif, J., & Susman, S. (1993). Demographic, psychosocial, and behavioral differences in samples of actively and passively consented adolescents. *Addictive Behaviors*, 18, 51-56.

DeVellis, R. F. (2003). *Scale development: Theories and applications* (2nd ed.). Thousand Oaks, CA: Sage.

Dionne, R. R. (1992). *Effective strategies for handling teasing among fifth and sixth grade children*. Unpublished doctoral dissertadon, California School of Professional Psychology, San Diego, CA.

Dumas, M. D., Dionne, R. R., Foster, S. L., Chang, M. K., Achar, M. S., & Martinez, C. M., Jr. (1992, November). *A critical incidents analysis of strategies for handling peer provocation among second-, fifth-, and eighth-grade children.* Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy, Boston.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research.* Hinsdale, IL: Dryden Press.

Edwards, A. L. (1970). *The measurement of traits by scales and inventories.* New York: Holt, Rinehart & Winston.

Edwards, A. L. (1990). Construct validity and social desirabiliy. *American Psychologist*, 45, 287-289.

Einstein. A. (1974). *The universe and Dr. Einstein.* New York: Lincoln Barnett.

Ellis, A., & Harper, R. A. (1961). *A guide to rational living.* North Hollywood, CA: Wilshire.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-299.

Farrell, A. D. (1999). Statistical methods in clinical research. In P. C. Kendall, J. N. Butcher, & G. N. Holmbeck (Eds.), *Handbook of research methods in clinical psychology* (2nd ed., pp. 72-106). New York: Wiley.

Ferrna, E. K. (2005). *Peer influences on body image related concerns: The role of social comparison.* Unpublished doctoral dissertation, Alliant International University, San Diego, CA.

Finch, C. L. (2001). *The relationship among relationally aggressive behavior, emotion and social cognitions in preadolescent females.* Unpublished doctoral dissertation, Alliant International University, San Diego, CA.

Fine, M. A., & Kurdek, L A. (1993). Reflections on determining authorship credit and anthorship order on faculty-student collaborations. *American Psychologist*, 48, 1141-1147.

Fink, A. (2002). *The survey handbook* (2nd ed.). Thousand Oaks, CA: Sage.

Fischer, J., & Corcoran, K. J. (2000a). *Measures for clinical practice: Adults.* New York: Free Press.

Fischcr, J., & Corcoran, K. J. (2000b). *Measures for clinical practice: Couples, families, and children.* New York: Free Press.

Fisher, C. B., Hoagwood, K. Boyce, C., Duster, T., Frank, D. A., Grisso, T., et al. (2002). Research ethics for mental health science involving ethnic minority children and youth. *American Psychologist*, 57, 1024-1040.

Floyd, F., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 42, 286-299.

Foster, S. L., Bell-Dolan, D. J., & Burge, D. A. (1988). Behavioral observation. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (3rd ed., pp. 119-160). Elmsford, NY: Pergamon Press.

Foster, S. L., & Cone, J. D. (1995). Validity issues in clinical assessment. *Psychological Assessment*, 7, 248-260.

Foster, S. L., Laverty Finch, C., Gizzo, D., & Osantowski, J. (1999). Practical issues in self-observation. *Psychological Assessment*, 11, 426-438.

Foster, S. L., & Martinez, C. R., Jr. (1995). Ethnicity: Conceptual and methodological issues in child clinical research. *Journal of Clinical Child Psychology*, 24, 214-226.

Frick, R. W. (1996). The appropriate use of null hypothesis testing. *Psychological Methods*, 1, 379-390.

Cantt charts. (2005). Retrieved November 17, 2005, from http://www.ganttcharts.com/index.html

Galvan, J. L. (1999). *Writing literature reviews.* Los Angeles: Pyrczak.

Galvan, J. L. (2004). *Writing literature reviews: A guide for students of the social and behavioral sciences* (2nd ed.). Los Angeles: Pyrczak.

Gelfand, H. G., & Walker, C. J. (Eds.). (2001). *Mastering APA style: Student's workbook and training guide.* Washington, DC: American Psychological Association.

Gevirtz, R. (2002). Physiological perspectives. In B. Horwitz (Ed.), *Communication apprehension: Origins and management* (pp. 114-136). Florence, KY: ThompsonDelmar Learning.

Goldman, B. A., & Mitchell, D. F. (Eds). (2003). *Directory of unpublished experimental mental measures* (Vol. 8). Washington, DC: American Psychological Association.

Gorsuch, R L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Goyer, L. E. (2005). *Responsiveness to deviancy training among highly deviant, moderately deviant, and non-deviant early adolescent males.* Unpublished doctoral dissertation, Alliant International University, San Diego, CA.

Green, S. B., & Salkind, N. J, (2004). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

Greenberger, D., & Padesky, C. (1995). *Mind over mood: Change how you feel by changing the way you think.* New York: Guilford Press.

Grimm, L. G., & Yarnold, P. R. (Eds.). (1995). *Reading and understanding multivariate statistics.* Washington, DC: American Psychological Association.

Grimm, L. G., & Yarnold, P. R. (Eds.). (2000). *Reading and understanding MORE multivariate statistics.* Washington, DC: American Psychological Association.

Grusec, J. (1991). The socialization of altruism. In M. S. Clark (Ed.), *Prosocial behavior* (pp. 9-33). Newbury Park, CA: Sage.

Guilford, J. P. (1956). *Fundamental statistics in psychology and education.* New York: McGraw-Hill.

Hartmann, D. P. (1982). Assessing the dependability of di-

rect observational data. In D. P. Hartmann (Ed.), *New directions for methodogy of social and behavioral science: Using observers to study behavior* (pp. 51-65). San Francisco: Jossey-Bass.

Hayes, S. C., Barlow, D. H., & Nelson-Gray, R. O. (1999). *The scientist practitioner: Research and accountability in the age of managed care* (2nd ed.). Boston: Allyn & Bacon.

Haynes, S. N., Richard, D. C. S., & Kubany, E. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 42, 238-247.

Helmes, E., & Holden, R. (2003). The construct of social desirability: One or two dimensions. *Personality and Individual Differences*, 34, 1015-1023.

Henry, G. T. (1994). *Graphing data*. Thousand Oaks, CA: Sage.

Heppner, P. P., & Heppner, M. J. (2004). *Writing and publishing your thesis, dissertation, and research*. Belmont, CA: Thompson-Brooks/Cole.

Hersen, M., & Bellack, A. S. (Eds.). (1988). *Dictionary of behavioral assessment techniques*. Elmsford, NY: Pergamon Press.

Hochhauser, M. (1999). Informed consent and patients' rights documents: A right, a rite, or a rewrite? *Ethics and Behavior*, 9, 1-20.

Hoier. T. S. (1984). *Target selection of social skills for children: An experimental investigation of the template matching procedure*. Unpublished doctoral dissertation, West Virginia University, Morgantown.

Hoier, T. S., & Cone, J. D. (1987). Target selection of social skills for children: The template matching procedure. *Behavior Modification*, 11, 137-163.

Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, 65, 599-610.

Horowitz, M. J., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: A measure of psychosomatic stress. *Psychosomatic Medicine*, 41, 209-218.

House, A. E., House, B. J., & Campbell, M. D. (1981). Measures of interobserver agreement: Calculation formulas and distribution effects. *Journal of Behavioral Assessment*, 3, 37-57.

Huberty, C. J., & Morris, J. D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin*, 105, 302-308.

Huitema, B. E. (1980). *The analysis of covariance and alternatives*. New York: Wiley.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis* (2nd ed.). Newbury Park, CA: Sage.

Inderbitzen-Pisaruk, H., & Foster, S. L. (1990). Adolescent friendships and peer acceptance: Implications for social skills training. *Clinical Psychology Review*, 10, 425-439.

Jaccard, J., Becker, M. A., & Wood, G. (1984). Pairwise multiple comparison procedures: A review. *Pyscholoical Bulletin*, 96, 589-596.

Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and tactics of human behavioral research* (2nd ed). Hillsdale, NJ: Erlbaum.

Kaplan, R. M., & Saccuzzo, D. P. (2005). *Psychological testing: Principles, applications, and issues* (6th ed.). Belmont, CA: Wadsworth.

Kazdin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Boston: Allyn & Bacon.

Kelley, M. L., Reitman, D., & Noell, G. H. (2002). *Practitioner's guide to empirically based measures of school behavior*. New York: Kluwer.

Kelley, T. L. (1927). *Interpretation of educational measurements*. Yonkers-on-Hudson, NY: World Book.

Kent, R. N., O'Leary, K. D., Diament, C., & Dietz, A. (1974). Expectation biases in observational evaluation of therapeutic change. *Journal of Consulting and Clinical Psychology*, 42, 774-780.

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook*. (4th ed.). Upper Saddle River, NJ: Pearson Prentice-Hall.

Kerlinger, F. N. (2000). *Foundations of behavioral research*. (4th ed.). New York: Harcourt Brace.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole Wadsworth.

Klem, L. (1995). Path analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 65-97). Washington, DC: American Psychological Association.

Koyre, A. (1965). *Newtonian studies*. Cambridge, MA: Harvard University Press.

Krathwohl, D. R. (1988). *How to prepare a research proposal: Suggestions for those seeking funds for behavioral science research* (3rd ed.). Syracuse, NY: School of Education, Syracuse University.

Kuhn, T. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Lakin, A. (1974). *How to get control of your time and your life*. New York: Signet Books.

Lang, P. J. (1971). The application of psychophysiological methods to the study of psychotherapy and behavior modification. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 75-125). New York: Wiley.

Licht, M. H. (1995). Multiple regression and correlation. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 19-64). Washington, DC: American Psychological Association.

Lipsey, M. W., & Wilson, D. B. (2000). *Practical meta analysis*. Thousand Oaks, CA: Sage.

Lord, C. (2005). *Indirect aggression and social support among elderly retirement community residents.* Unpublished doctoral dissertation, Alliant International University, San Diego, CA.

MacCallum, R. C., Widamon, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4, 84-99.

Mackinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7, 83-104.

Maher, B. A. (1978). A reader's, writer's, and reviewer's guide to assessing research reports in clinical psychology. *Journal of Consulting and Clinical Psychology*, 46, 835-838.

Mahoney, M. J., & Mahoney, B. K. (1976). *Permanent weight control.* New York: Norton.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison approach* (2nd ed.). Mahwah, NJ: Erlbaum.

McBurney, D. H. (1990). *Experimental psychology* (2nd ed.). Belmont, CA: Wadsworth.

McBurney, D. H. (2004). *Research methods* (6th ed.). Belmont, CA: Brooks/Cole.

McInerney, D. M. (2001). *Publishing your psychology research: A guide to writing for journals in psychology and related fields.* Thousand Oaks, CA: Sage.

Meltzoff, J. (1998). *Critical thinking about research: Psychology and related fields.* Washington, DC: American Psychological Association.

Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods* (3rd ed.). Glendale, CA: Pyrczak.

Merriam-Webster OnLine. (n. d.). Retrieved March 7, 2005, from http://www.m-w.com/cgibin/dictionary?book=Dictionary&va=dissertation&x=19&y=16m

Miller, D. C. (2002). *Handbook of research design and social measurement* (6th ed.). Newbury Park, CA: Sage.

Murphy, K. R., & Myors, B. (2004). *Statistical power analysis* (2nd ed.). Mahwah, NJ: Erlbaum.

Murphy, L. L., Plake, B. S., Impara, J. C., & Spies, R. A. (Eds.). (2002). *Tests in print VI.* Lincoln: University of Nebraska.

Myers, J. L., DiCecco, J. V., White, J. B., & Borden, V. M. (1982). Repeated measurements on dichotomous variables: *Q* and *F* tests. *Psychological Bulletin*, 92, 517-525.

Nagy, T. F. (2005). *Ethics in plain English: An illustrative casebook for psychologists* (2nd ed.). Washington, DC: American Psychological Association.

National Institutes of Health. (2005a). *Guidelines for the conduct of research involving human subjects at the National Institutes of Health.* Retrieved November 7, 2005, from http://ohsr.od.nih.gov/guidelines/graybook.html

National Institutes of Health. (2005b). *Regulations and ethical guidelines.* Retrieved November 7, 2005, from http://ohsr.od.nih.gov/guidelines/45cfr46.html

Nelson, G. L., Cone, J. D., & Hanson, C. R. (1975). Training correct utensil use in retarded children: Modeling vs. physical guidance. *American Journal of Mental Deficiency*, 80, 114-122.

Neuman, W. L. (1997). *Social research methods: Qualitative and quantitative approaches* (3rd ed.). Boston: Allyn & Bacon.

Newton, R. R., & Rudestam, K. E. (1999). *Your statistical consultant: Answers to your data analysis questions.* Thousand Oaks, CA: Sage.

Nezu, A., Ronan, G. F., Meadows, E. A., & McClure, K. S. (2000). *Practitioner's guide to empirically based measures of depression.* New York: Kluwer.

Nicol, A. A. M., & Pexman, P. M. (1999). *Presenting your findings: A practical guide for creating tables.* Washington, DC: American Psychological Association.

Nunnally, N. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Oetting, E. R. (1986). Ten fatal mistakes in grant writing. *Professional Psychology: Research and Practice*, 17, 570-573.

O'Leary, K D., Kent, R. N., & Kanowitz, J. (1975). Shaping data congruent with experimental hypotheses. *Journal of Applied Behavior Analysis*, 8, 463-469.

O'Shea, H. (2003). *Assessment of aggression in children: An exploration of self and peer report methods.* Unpublished dissertation proposal, Alliant International University, San Diego, CA.

Pallant, J. (2001). *SPSS survival manual: A step by step guide to data analysis using SPSS.* Philadelphia: Open University Press.

Parsonson, B. S., & Baer, D. M. (1978). The analysis and presentation of graphic data. In T. R. Kratochwill (Ed.), *Single subject research* (pp. 101-165). San Diego, CA: Academic Press.

Paul, G. L. (1966). *Insight vs. desensitization in psychotherapy.* Stanford, CA: Stanford University Press.

Paul, G. L. (1969). Behavior modification research: Design and tactics. In C. M. Franks (Ed.), *Behavior therapy: Appraisal and status* (pp. 29-62). New York: McGraw-Hill.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N., Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49-69). Mahwah, NJ: Erlbaum.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Orlando,

FL:Harcourt Brace.

Pedhazur,E. J. ,& Schmelkin,L. P. (1991). *Measurement, design, and analysis: An integrated approach.* Hillsdale, NJ:Erlbaum.

Plagiarism: What it is and how to avoid it. (n. d.). Retrieved February 4,2005, from http://www. indiana. edu/ ~ wts/pamphlets. html

Plake, B. S. , Impara, J. C. , & Spies, R. A. (Eds.). (2003). *The fifteenth mental measurements yearbook.* Lincoln:University of Nebraska.

Preacher, K. J. , & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics*,2,13-43.

Pyrczak, F. (2000). *Completing your thesis or dissertation.* Los Angeles:Author.

Raudenbush,S. W. ,& Bryk,A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*(2nd ed.). Thousand Oaks,CA:Sage.

Ray,W. J. (2006). *Methods: Toward a science of behavior and experience* (8th ed.). Pacific Grove, CA: Brooks/Cole.

Raykow,T. ,& Marcoulides,G. M. (2000). *A first course in structural equation modeling.* Mahwah,NJ:Erlbaum.

Research and Education Association. (1981). *Handbook of psychiatric rating scales.* New York:Author.

Roberts,R. E. ,Lewinsohn,P. M. ,& Seeley,J. R. (1991). Screening for adolescent depression: A comparison of depression scales. *Journal of the American Academy of Child and Adolescent Psychiatry*,30,58-66.

Robinson,J. P. ,Shaver,P. R. ,& Wrightsman,L. S. (Eds.). (1991). *Measures of personality and social psychological attributes.* San Diego,CA:Academic Press.

Rodgers,W. (1995). Analysis of cross-classified data. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics*(pp. 169-216). Washington,DC:American Psychological Association.

Romanczyk,R. G. ,Kent,R. N. ,Diament,C. ,& O'Leary, K. D. (1973). Measuring the reliability of observational data: A reactive process. *Journal of Applied Behavior Analysis*,6,175-184.

Rosenthal,R. (1969). Interpersonal expectations: Effects of the experimenter's hypothesis. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 181-277). San Diego,CA:Academic Press.

Rosenthal,R. (1991). *Meta-analytic procedures for social research.* Newbury Park,CA:Sage.

Rosenthal,R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*,118,183-192.

Rossi,J. S. (1990). Statistical power of psychological research: What have we gained in 20 years? *Journal of Consulting and Clinical Psychology*,58,646-656.

Rudestam,K. E. ,& Newton,R. R. (2001). *Surviving your dissertation*(2nd ed.). Thousand Oaks,CA:Sage.

Sales,B. D. ,& Folkman,S. (Eds.). (2000). *Ethics in research with human participants.* Washington,DC:American Psychological Association.

Sanchez-Hucles,J. ,& Cash,T. F. (1992). The dissertation in professional psychology programs: 1. A survey of clinical directors on requirements and practices. *Professional Psychology: Research and Practice*,23. 59-61.

Sandler,R. B. (2000). *Plagiarism in colleges in the USA.* Retrieved February 2,2005,from http://www. apa. org/

Saving your thesis when support goes down the drain. (1995). *Dissertation News*,3,1-7.

Schafer,J. L,& Graham,J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*,7, 147-177.

Schwarz,N. ,& Oyserman,D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation*,22, 127-160.

Severino,J. (2002). *Gay and straight men's attributions to simulated antigay and nonbias victimization.* Unpublished dissertation proposal,Alliant International University,San Diego,CA.

Shadish,W. R. ,Cook,T. D. ,& Campbell,D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston:Houghton Mifflin.

Shaw,R. (2002). Grammar checkers: Helpful or harmful? *USA Today.* Retrieved November 17,2005,from http://www. usatoday. com/tech/news/2002/07/15/tech-grammar-full. htm

Shawchuck,C. R. ,Fatis,M. ,& Breitenstein,J. L. (1986). A practical guide to the assignment of authorship credit. *Behavior Therapist*,9,216-217.

Shiffman, S. (2000). Real-time self-report of momentary states in the natural environment: Computerized ecological momentary assessment. In A. A. Stone,J. S. Turkan, C. A. Bachrach,J. B. Jobe,H. S. Kurtzman,& V. S. Cain (Eds.), *The science of self-report: Implications for research and practice* (pp. 277-296). Mahwah, NJ: Erlbaum.

Shiller,R. J. (1999). Human behavior and the efficiency of financial markets. In J. B Taylor & M. Woodford (Eds.), *Handbook of Macroeconomics* (Vol. 1, pp. 1305-1340). Amsterdam:Elsevier.

Shuller,D. Y. ,& McNamara,J. R. (1976). Expectancy factors in behavioral observation. *Behavior Therapy*, 7, 519-527.

Siegel,S. (1956). *Nonparametric statistics for the behavioral sciences.* New York:McGraw-Hill.

Sikora,D. M. (1989). *Divorce, environmental change, parental conflict, and the peer relations of preschool children.* Unpublished doctoral dissertation,West Virginia University,Morgantown.

Singer,J. D. ,& Willett,J. S. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York:Oxford University Press.

Smith, J. C. (2005). *Relaxation, meditation, and mindfulness: Self-training manual.* Morrisville NC: Lulu Press.

Smyth, J. M., & Stone, A. A. (2003). Ecological momentary assessment research in behavioral medicine. *Journal of Happiness Studies*, 4, 35-52.

Sobell, M. B., Bogardis, J., Schuller, R., Leo, G. I., & Sobell, L. C. (1989). Is self-monitoring of alcohol consumption reactive? *Behavioral Assessment*, 11, 447-458.

Solso, R. L., & MacLin, M. (2002). *Experimental psychology: A case approach* (7th ed.). Boston: Allyn & Bacon.

Sternberg, D. (1981). *How to complete and survive a doctoral dissertation.* New York: St. Martin's Press.

Stone, A. A., Kessler, R. C., & Haythornthwaite, J. A. (1991). Measuring daily events and experiences: Decisions for the researcher. *Journal of Personality*, 59, 575-607.

Stone, A. A., Turkan, J. S., Bachrach, C. A., Jobe, J. B., Kurtz man, H. S., & Cain. V. S (Eds.). (2000). *The science of self-report: Implications for research and practice.* Mahwah, NJ: Erlbaum.

Stuart, R. B. (Ed.). (1977). *Behavioral self-management: Strategies, techniques and outcome.* New York: Brunner/Mazel.

Suen, H. K., & Ary, D. (1989). *Analyzing quantitative behavioral observation data.* Hillsdale, NJ: Erlbaum.

Sweetland, R. C., & Keyser, D. J. (Eds.). (1983). *Tests: A comprehensive reference for assessments in psychology, education and business.* Kansas City, MO: Test Corporation of America.

Sweetland, R. C., & Keyser, D. J. (Eds.). (1997). *Tests* (4th ed.). Austin, TX: PRO-ED.

Tabachnick, B. F., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Taplin, P. S., & Reid, J. B. (1973). Effects of instructional set and experimental influence on observer reliability. *Child Development*, 44, 547-554.

Thomas, L., & Krebs, C. J. (1997). A review of statistical power analysis software. *Bulletin of the Ecological Society of America*, 78, 126-139.

Thompson, B. (2000). Ten commandments of structural equation modeling. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding MORE multivariate statistics* (pp. 261-283). Washington, DC: American Psychological Association.

Thompson, B. (Ed.). (2003). *Score reliability: Contemporary thinking on reliability issues.* Newbury Park, CA: Sage.

Toastmaster's International. (2005). Retrieved November 17, 2005, from http://www. toastmasters. org/

Tracy, B. (2004). *Time power: A proven system for getting more done in less time than you ever thought possible.* New York: AMACOM.

Tryon, W. W. (1998). Behavioral observation. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (4th ed., pp. 79-103). Needham Heights, MA: Allyn & Bacon.

Tufte, E. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, CT: Graphics Press.

UCLA Academic Technology Services. (n. d.) *Statistical computing resources.* Retrieved November 15, 2005, from http://www. ats. ucla. edu/stat

Ullman, J. B. (2001). Structural equation modeling. In B. F. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (4th ed., pp. 653-771). Needham Heights, MA: Allyn & Bacon.

University of Washington School of Public Health. (2005). *Creating a poster using MS PowerPoint.* Retrieved November 18, 2005, from http://depts. washington. edu/mphpract/pp poster. html

Vernoff, J. (2001). Writing. In K. E. Rudestam & R. R. Newton (Eds.), *Surviving your dissertation* (2nd ed.). Thousand Oaks, CA: Sage.

von Eye, A., & Mun, E. Y. (2005). *Analyzing rater agreement: Manifest variable methods.* Mahwah, NJ: Erlbaum.

Wallgren, A., Wallgren, B., Persson, R., Jorner, U., & Haaland, J. (1996). *Graphing statistics & data.* Thousand Oaks, CA: Sage.

Webb, E. J. (1961). The choice of problem. *American Psychologist*, 16, 223-227.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally.

Weisberg, H. F., Krosnick, J. A., & Bowen, B. D. (1996). *An introduction to survey research, polling, and data analysis* (3rd ed.). Newbury Park, CA: Sage.

Whalen, C. K., Jamner, L. D., Honker, B., & Delfino, R. J. (2001). Smoking and moods in adolescents with depressive and aggressive dispositions: Evidence from surveys and electronic diaries. *Heath Psychology*, 20, 99-111.

Wiggins, J. S. (1959). Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *Journal of Consulting Psychology*, 23, 419-427.

Wilkinson, L., & The Task Force on Statistical Inference. (1999). Statistical methods in psychology journals. *American Psychologist*, 54, 594-604.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.

Wolfe, V. V. (1986). *Paternal and marital factors related to child conduct problems.* Unpublished doctoral dissertation, West Virginia University, Morgantown.

Wolfe, V. V., Gentile, C., Michienzi, T., Sas, L., & Wolfe, D. A. (1991). The Children's Impact of Traumatic Events Scale: A measure of post-sexual-abuse PTSD symptoms. *Behavioral Assessment*, 13, 359-383.

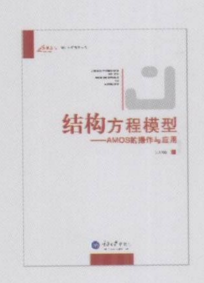
Yates, B. T. (1982). *Doing the dissertation: The nuts and bolts of psychological research.* Springfield, IL: Charles C. Thomas.

策划人：雷少波　林佳木



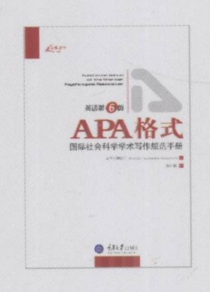
《量化研究与统计分析——
SPSS中文视窗版数据分析范例解析》

本书立足于为量化研究的学习者服务，分别从量化研究的基本思路和方法、统计原理和技术、软件应用三部分进行材料的有机组织，并对二十余种常用量化分析技术辅以完整的范例，以帮助读者在整体上学会做量化研究。本书适合社会科学各学科高年级本科生、硕博士研究生自学，也可作为相关课程的教材。



《结构方程模型——
AMOS的操作与应用》

本书详细讲解和演示结构方程模型多种分析方法的操作步骤，是一本理想的AMOS与结构方程模型应用方面的指导读物。本书的“使用者界面”取向，使得不懂传统SEM语法的使用者也能在最短时间内学会用AMOS绘制各种SEM模型图，并将模型估计、模型识别判断、模型修正与模型验证，实际应用于自己的研究领域中。



《APA格式——国际社会科学学术写作手册》(第6版)

这是一部名副其实的世纪经典，经过近100年的发展（1928年首次公开出版），它的影响力已经远远超出了心理学的领域，成为国际社会科学界普遍认可的最重要的学术论文写作规范。



《心理学研究要义》

每年都要指导各种层次的学生做本科生基金研究或毕业论文，每次都要强调书中涉及的这些内容。常常在实验室的例会上说，这次说过了，就不能再出现这类问题啦！其实不可能的。原则和规范是定的，而学生是常新的。所以如果这本书能够作为一本很实用的简明指导手册，在学生进行自己的研究工作时能时不时地翻看一下，以减少迷茫和困惑，少走很多弯路的话，我们的目的就达到了。